

**Università degli Studi di Padova**

---

DIPARTIMENTO DI MATEMATICA

Corso di Laurea in Matematica

TESI DI LAUREA MAGISTRALE

## **Reconstruction of medical images from Radon data in transmission and emission tomography**

Candidato:

**Davide Poggiali**

Relatore:

**Prof. Stefano De Marchi**

Correlatore:

**Dr. Diego Cecchin**



### **Ringraziamenti**

Desidero ringraziare chi ha sostenuto e incoraggiato il proseguimento dei miei studi fino a questo punto, *in primis* la mia famiglia, sempre critica e affettuosa nei miei confronti, Silvia e la sua famiglia, gli amici tutti.

Ringrazio il prof. De Marchi e il prof. Cecchin per avermi seguito passo dopo passo in questo lavoro svolto su due fronti. Ringrazio inoltre le Università di Firenze e Padova e tutte le persone che con il loro lavoro hanno permesso e aiutato il mio studio.



## Essential glossary

- **analytical:** precise, exact. Analytical methods are based on analytical inversion formulas.
- **attenuation:** physical phenomenon of the deviation of an **electromagnetic ray** from its natural path, due to the contact of a body.
- **CT:** X-ray Computed Tomography, a medical imaging procedure that utilizes computer-processed **X-rays** to produce 'slices' of specific areas of the body.
- **dummy:** (also called **phantom**, we use that name to avoid confusions) physical device that simulates a certain organ and used to make experiments on the tomography machine, *in vitro*.
- **electromagnetic waves:** is an oscillation which propagates in the space, transporting energy but not matter. We distinguish different kind of waves by their wavelength, among which:
  - **X-rays:** wavelength in the range of 0.01 to 10 nanometers.
  - **Gamma-rays:** wavelength shorter than 1 picometer.
- **ill-conditioned problem:** given  $f(x) = y$ ,  $y \in Y$ , we have to find  $x \in X$ , with  $X$  and  $Y$  vector spaces. This problem is i.c. if an arbitrarily small perturbation on  $y$  can lead to an arbitrarily large error on  $x$ .
- **in silico:** experiment made on mathematical simulation of the data (in this work we use **phantoms**) and not using real data.
- **in vitro:** experiments made on real data coming from a biological/chemical phenomenon reproduced in a test tube (in this work we use a **dummies**), and not using an organism.
- **kernel function:** a function  $K$  such that  $K(x, y) = \phi(x - y)$  with  $\phi$  a **RBF**.
- **nuclear medicine:** a branch of medicine involving the application of radioactive substances in the diagnosis and treatment of disease.

- **over(/under)-determined:** a problem with more(/less) equations than unknowns. Given  $f(x) = y$ ,  $y \in Y$ , we have to find  $x \in X$ , with  $X$  and  $Y$  vector spaces. This problem is o.d. if  $\dim(Y) > \dim(X)$  and u.d. if  $\dim(Y) < \dim(X)$ .
- **phantom:** an image created to test the efficiency of the algorithms *in silico*, without need of data given by a physical machine. The most used phantom in this work is the **Shepp-Logan phantom**, which is similar to a brain, see Figure 3.2.
- **Radon transform:** is the integral transform consisting of the integral of a function over straight lines, proposed by the Austrian mathematician Johann Radon in 1917. It simulates mathematically the effect of **attenuation** of some **electromagnetic waves** passing through different materials at several angles, and then it is useful for the CT reconstruction.
- **RBF:** Radial Basis Function, functions  $\phi : \Omega \subseteq \mathbb{R}^s \rightarrow \mathbb{R}$  which depends only on the radial component  $r = \|x\|$ .
- **sampling:** times or points at which the data is collected from a continuous phenomenon.
- **scattering:** erroneous acquisition of some rays made by the tomography machine due to **attenuation**.
- **sinogram:** an  $l \times p$  matrix of data collected by the tomography machine in  $l$  angular scans. In the case of *in silico* experiments we call s. the discrete **Radon transform** of the **phantom**.
- **SPECT:** Single-Photon Emission Computed Tomography, is a **nuclear medicine** imaging technique using radioactive tracers that emit **Gamma-rays** in decaying.

# Introduction

The decorative initial 'T' is intricately designed with floral and foliate motifs, featuring a large, bold letter 'T' with delicate leaves and vines extending from its top and sides.

from above.

he problem of image reconstruction from projections dates back to the 1970s with the first X-ray machines, which give out as data the result of some circular acquisitions. An algorithm provides to extract the relevant section of the body, as if seen from above. We discuss about these algorithms and their efficiency both in the case of the transmission tomography (such as the well-known CAT, which emits and detects X-rays), and of the emission tomography (for instance PET, which detects pairs of positrons and SPECT, which captures gamma-rays) used in **nuclear medicine**<sup>1</sup>. In the first case a series of rays pass through the section before arriving at a detector in order to obtain a section of the organs, while in the second case the rays are coming from inside the patient's body, in which they were previously injected tracers (not to be confused with the radiocontrast agents) in order to obtain cross-sectional areas in which is concentrated and the tracer, e.g. an area affected by tumor or a working area of a certain organ.

In particular, we will discuss the hybrid SPECT/CT case, in which the SPECT image, obtained with a radioactive tracer with low emissions, is related to a low-resolution CT scan on the same section, in order to understand the relation between the area highlighted by the marker and the neighboring organs.

---

<sup>1</sup>nuclear medicine uses radioactive tracers, i.e. chemical elements with low radioactivity which are injected into the patient and which the human body uses as if they were normal atoms from physiological reactions. E.g. the  $^{18}FdG$ , 18-Flu-deoxy-glucose used in PET, which the body absorbs as if it were glucose, of which the tumor cells are “greedy”. From the annihilation or from the decay of these tracers is possible to recover their position.

# Contents

<b>Introduction</b>	<b>iii</b>
<b>1 Some medical and technical issues</b>	<b>1</b>
1.1 Electromagnetic waves . . . . .	1
1.2 Attenuation . . . . .	2
1.3 Scattering . . . . .	3
1.4 Partial volume effect . . . . .	5
1.5 Structure of a Gamma camera . . . . .	6
1.6 Resolution and sensitivity . . . . .	8
1.6.1 Septal thickness . . . . .	8
1.6.2 Geometric resolution . . . . .	9
1.6.3 Experimental resolution . . . . .	9
1.6.4 Sensitivity . . . . .	10
1.7 Tracers used in SPECT and standard exams . . . . .	11
1.8 Some assumptions . . . . .	12
<b>2 Mathematical modeling: the Radon transform and its variants</b>	<b>15</b>
2.1 The Beer-Lambert's law and the Radon transform . . . . .	15
2.2 Alternative Transforms . . . . .	16
2.2.1 Backprojection . . . . .	16
2.2.2 Divergent beam transform . . . . .	17
2.2.3 Attenuated Radon transform . . . . .	17
<b>3 Analytical properties and inversion of the Radon transform</b>	<b>19</b>
3.1 Basic properties . . . . .	20
3.2 Sampling and resolution . . . . .	21
3.3 Inversion formulas for the Radon transform . . . . .	24
3.4 Discretization of the FBP formula and error bound . . . . .	25

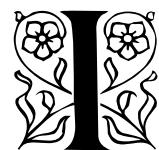
<b>4 Algebraic Reconstruction Techniques</b>	<b>31</b>
4.1 Some iterative methods . . . . .	32
4.1.1 The Kaczmarz algorithm . . . . .	33
4.1.2 Maximum Likelihood Expectation Maximization . . . . .	34
4.1.3 Conjugate gradient . . . . .	36
4.2 Accelerated reconstruction . . . . .	37
<b>5 Kernel methods</b>	<b>39</b>
5.1 Direct kernel interpolation . . . . .	39
5.2 Definite positivity . . . . .	41
5.3 Error estimate . . . . .	42
<b>6 The emission tomography case</b>	<b>47</b>
6.1 Geometric methods for attenuation correction . . . . .	47
6.1.1 Soreson algorithm . . . . .	48
6.1.2 Chang algorithm . . . . .	48
6.2 Inversion formula of the attenuated Radon transform . . . . .	49
6.3 Discretization of the AtRT inversion formula . . . . .	50
6.4 Iterative methods . . . . .	51
<b>7 Regularization methods and post-processing</b>	<b>55</b>
7.1 Low-pass filters . . . . .	55
7.2 Philips-Tichonov regularization . . . . .	56
7.3 A segmentation algorithm . . . . .	57
7.4 A simple algorithm for the contour detection of the ROI . . . . .	59
<b>8 Experimental results from <i>in silico</i> tests</b>	<b>61</b>
8.1 Filtered Back Projection results . . . . .	61
8.2 Iterative Methods results . . . . .	64
8.3 Kernel methods results . . . . .	67
8.4 Comparisons of the results . . . . .	71
8.5 Hybrid SPECT/CT simulations . . . . .	74
<b>9 Results from <i>In vitro</i> experiments</b>	<b>77</b>
9.1 Resolution and sampling . . . . .	77
9.2 Three capillaries experiment and experimental resolution . . . . .	79
9.3 Cerebral dummy . . . . .	82
<b>Conclusions and future work</b>	<b>85</b>
<b>Bibliography</b> . . . . .	<b>87</b>

# Chapter 1

## Some medical and technical issues

Mais vous, mon cher Pangloss, dit Candide, comment se peut-il que je vous revoie? Il est vrai, dit Pangloss, que vous m'avez vu pendre; je devais naturellement être brûlé mais vous vous souvenez qu'il plut à verse lorsqu'on allait me cuire: l'orage fut si violent qu'on désespéra d'allumer le feu: je fus pendu, parcequ'on ne put mieux faire: un chirurgien acheta mon corps, m'emporta chez lui, et me disséqua. Il me fit d'abord une incision cruciale depuis le nombril jusqu'à la clavicule. On ne pouvait pas avoir été plus mal pendu que je l'avais été. [...] Enfin je respirais encore: l'incision cruciale me fit jeter un si grand cri, que mon chirurgien tomba à la renverse.

Voltaire, *Candide*



In this chapter we will discuss some medical, physical and technical issues useful to better understanding the image reconstruction problem and to decide what assumptions we need in order to model the behavior of the machines with sufficient accuracy and thus to obtaining a good reconstruction.

### 1.1 Electromagnetic waves

An **electromagnetic wave** is an oscillation which propagates in the space, transporting energy but not matter. Its intuitive form is a cosine function with amplitude  $A$  and frequency  $\nu$ . Its speed in the vacuum is  $c \cong 3 \cdot 10^8 m/s$ , the speed of light, so if  $\lambda$  is the wavelength, then  $\lambda\nu = c$ . The energy carried

by the wave is  $E = h\nu$  (where  $h \cong 6 \cdot 10^{-34} Js$  is the Plank constant), hence the energy is inversely proportional to the wavelength:

$$E = \frac{hc}{\lambda} \quad (1.1)$$

Some examples of electromagnetic waves are the light of the sun, infrared, microwaves, radio waves, up to X-rays and gamma rays.

It is known as **X-rays** that portion of the electromagnetic spectrum with a wavelength in the range of  $10^{-3}$  up to  $10^1$  nm discovered in 1895 by the German physicist Wilhelm Conrad Röntgen, who earned for this discover the first Nobel price in 1901. X-rays are produced from the transition of electrons, and they can penetration through many objects, included the human body. For this reason X-rays are widely used in medicine to discover irregularities into the patient's body.

Similar to X-rays, but shorter in wavelength (then they transport more energy for the (1.1)), the **gamma rays** (sometimes written as  $\gamma$ -rays) are produced by a radioactive decaying<sup>1</sup>. As the X-rays, gamma rays can be dangerous and cause burns, cancer, and genetic alteration, so doctors must pay attention not to do useless or too hazardous exams, by limiting the energy of the waves used and the time of exposition.

## 1.2 Attenuation

Every electromagnetic waves passing thought a material mean is subject to an **attenuation** due to diffraction; in every point the ray has a probability  $p$  to deviate from its natural path<sup>2</sup>. The linear attenuation coefficient  $\mu$  is proportional to the probability of deviation while passing in  $1cm$  of matter;  $\mu$  depends on the material, on its density and on the energy carried by the wave. The linear attenuation coefficient is proportional to the density  $\rho$ , so the mass attenuation coefficient  $\mu_m = \frac{\mu}{\rho}$ , only depends on the energy and the material.

Table 1.1 shows the mass attenuation coefficients of different materials (water, sodium iodide and lead) and energies.

---

<sup>1</sup>this is the fundamental difference between X and gamma rays: the first ones come from outside the nucleus of the atoms, the second ones from inside.

<sup>2</sup>see [2] Ch 6D, p.80 for a complete and detailed survey about photon attenuation and scattering.

Photon energy (eV)	Mass attenuation coefficient $\mu_m$ ( $cm^2/g$ )		
	$H_2O$	$NaI$	$Pb$
80	0.179	3.00	2.07
100	0.168	1.64	5.23
150	0.149	0.590	1.89
200	0.136	0.314	0.945
300	0.118	0.158	0.383
400	0.106	0.112	0.220
500	0.097	0.092	0.154
600	0.089	0.080	0.120
800	0.079	0.066	0.085
1000	0.070	0.058	0.069

Table 1.1: Mass attenuation coefficient for several values of the photon energy and for different materials.

It is evident from this table that the mass attenuation coefficient is decreasing with the growth of the energy or with the decreasing of the density, but it is neither inversely proportional to the energy or proportional to the density. Another popular way to represent this attenuation is the *Hounsfield scale*

$$H = \frac{\mu - \mu_{water}}{\mu_{water}}$$

which compares the attenuation coefficient of the medium and that of the water.

Computerized tomography, in its classical form, calculates the attenuation coefficient of the different areas inside the body from the difference between entering and outgoing X-rays, in order to recognize the different organs and find possible malfunctions or malformations.

### 1.3 Scattering

During the reconstruction we have to face different kind of issues related to the attenuation. The most known, called **scattering**, is the effect of a multiple attenuation of the same ray. In fact assume that a ray is diffracted in the point  $(x_1, y_1)$  and deviates from its natural path; assume also that, as in Figure 1.1, it deviates again in the point  $(x_2, y_2)$  in the direction of the detector (perpendicular in the figure). This means that the machine will see many rays “out of place”, but with low energies.

The most common way to reduce the scattering effect<sup>3</sup> is a Hamming windowing of the incoming data around the known peak of energy of the used tracer or ray, as shown in Figure 1.2.

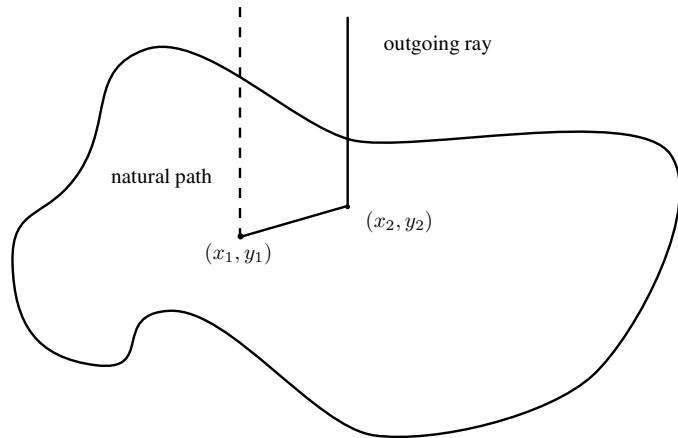


Figure 1.1: Scattering effect.

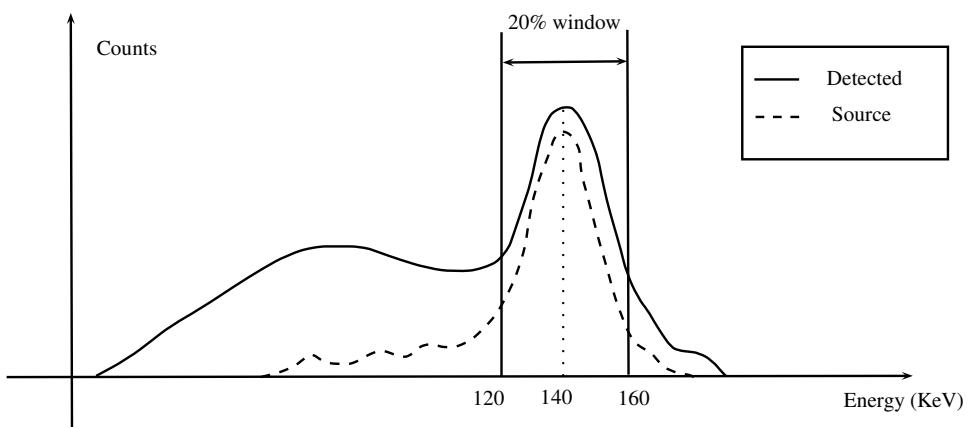


Figure 1.2: Scattering correction example for  $^{99m}Tc$ , who has a single peak of energy of 140KeV.

---

<sup>3</sup>See [2] Ch 10C p.161 for further explanations.

## 1.4 Partial volume effect

Another typical issue in tomography is the correct reconstruction of the higher-valued areas. In fact in such area the reconstruction will result smoother and lower-valuer than the original ones. Conversely the lower-valued areas next to it will have a growth of value near the higher area. This is known as **partial volume effect**, and it is similar to the light bulb effect: it is hard to see the exact position of the bulb since it is on because of the light rays next to it. As we will see later the partial volume effect is not related to the reconstruction algorithm we use, but only depends from resolution.

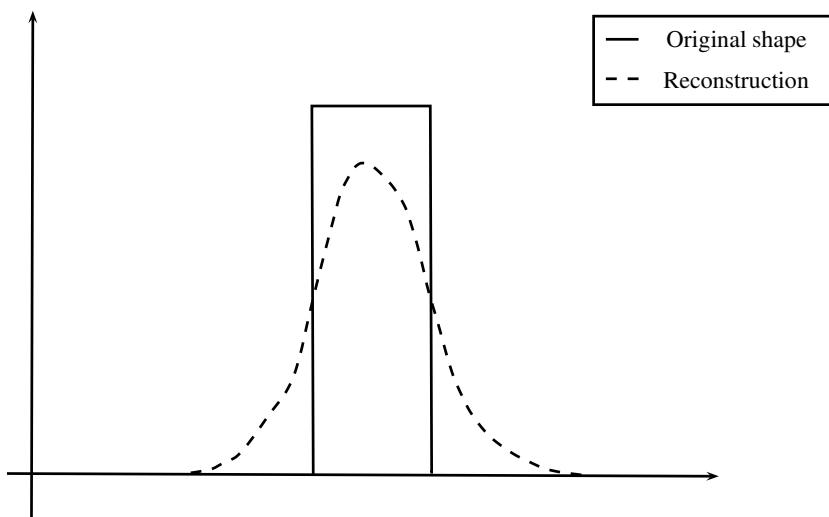


Figure 1.3: Partial volume effect example on the characteristic function of an interval.

## 1.5 Structure of a Gamma camera



Figure 1.4: External view of a Gamma-camera.

A machine for axial tomography externally has a toroidal structure. Inside of it is placed the patient through a sliding bed. The position of the couch obviously determines the height of the body of the patient at which we want to obtain the section. Some algorithms give a three-dimensional reconstruction from the data taken at different heights; other a two-dimensional reconstruction for each set of data relating to the same height. Inside the machine there are the emitters (only in the case of a transmission tomography) and the detectors.

Now, let us describe the structure of the detectors of the Gamma-camera

used in SPECT; the structure of other types of detectors are similar. The gamma rays leaving the highlighted area are selected by the **collimators**. The collimators are devices with circular, square or hexagonal section long  $L$ , and placed parallel to each other at a constant distance and normal to the circumference of capture. The hole has length  $D$  and the septal thickness (the space between two holes) is  $t$ . Their aim is to filter incoming rays passing only those nearly parallel (this is named *parallel-hole tomography*). Other dispositions of the collimators give different data acquisition (*divergent/convergent-hole or pin-hole tomography*).

The ray that passes out the collimators arrives in a scintillator crystal that

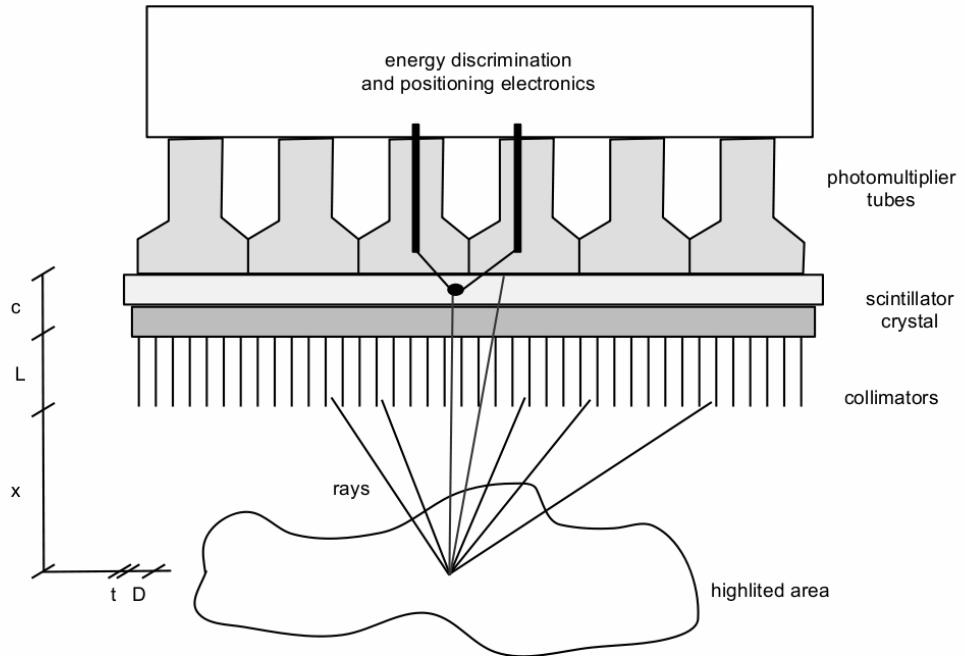


Figure 1.5: Structure of a Gamma-camera used in SPECT.

converts the beam into a spark of low-density. This spark gets in the photomultiplier tubes (or PM tubes), which converts the sparkle into variable beams of electrons with an energy of  $O(10^5)$  times greater than that of the sparkle. The machine will successively treat these beams as electrical pulses, convert to numerical data, and detect the location of the pulses; this last process is made by an algorithm which tries to indicate from which hole between the collimators had passed the original gamma ray.

Furthermore we call  $c$  the distance between the collimators and crystal,  $x$

the object-collimator distance and  $L_{eff} = L - \frac{2}{\mu}$  the effective length of the collimator with  $\mu$  attenuation coefficient of the material of the collimators<sup>4</sup>. From this knowledge on the functioning of the machine we can derive the sensitivity and the resolution of the instrument<sup>5</sup>. Once estimated the resolution and sensitivity, we also get the *a priori* error, which we will use during the experiments.

## 1.6 Resolution and sensitivity

### 1.6.1 Septal thickness

Does the collimator really stop all the non-parallel  $\gamma$ -rays? There is not any material able to stop all the rays, and, by consequence we have to set a level of septal penetration which can be considered negligible (e.g. a tolerance of 5 percent) and set the thickness as the corresponding value. Let  $w$  be the shortest path length for  $\gamma$ -rays to travel from one hole to the next as in Figure 1.6. Then, since  $t \ll l$ , we find the proportion  $t + 2D : l = t : w$ , which leads to the equality  $w = \frac{lt}{t+2D}$ .

Since the septal penetration has to be less than 5 percent, then for Beer-

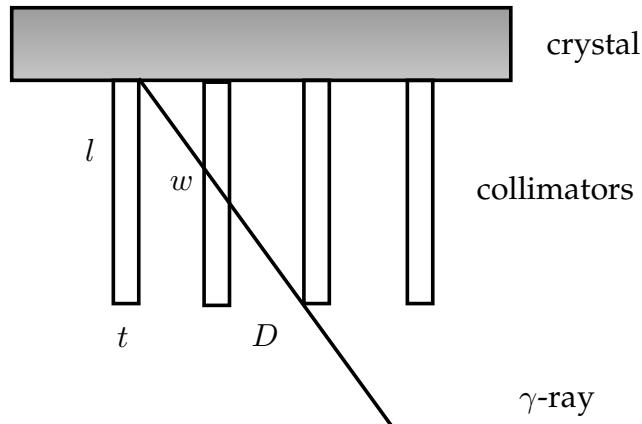


Figure 1.6: Scheme for the minimal septal thickness which gives the desired tolerance; note that  $t + 2D : l = t : w$ .

<sup>4</sup>usually lead,  $\mu = 2.49\text{mm}^{-1}$ .

<sup>5</sup>For further details on the following equations see [10] Ch 5, p.87.

Lambert's law<sup>6</sup>  $e^{\mu w} \leq 0.05$ ; we know that  $e^{-3} \cong 0.05$ , then  $w \gtrsim \frac{3}{\mu}$ , thus

$$t \gtrsim \frac{6D}{\mu l - 3}$$

This means that if we want to reduce the septal thickness the collimators must be made of material with high density and atomic number. Lead is the most used material because is cheap and easy to find; other materials, as tantalum, tungsten, and even gold have been employed in experiments<sup>7</sup>. From the formula above we also deduce that if the energies emitted by the source is bigger, then the attenuation coefficient will decrease, and the minimal septal thickness must be larger.

### 1.6.2 Geometric resolution

What is the smallest distance between two objects such that the machine can recognize them as two? To answer this question let us consider a point source of gamma rays as in Figure 1.5, and the two red rays in the figure (the parallel one and the last entering one). If  $R$  is the distance of the two rays at the PM tubes, then, by the similitude of triangles  $L_{eff} : D = (L_{eff} + x + c) : R$ , hence

$$R = D \left( 1 + \frac{x + c}{L_{eff}} \right)$$

since  $c \ll x$  the formula becomes

$$R \cong D + x \frac{D}{L_{eff}}$$

The resolution of the machine depends on the collimator resolution and on the intrinsic resolution (the resolution of the crystal and the electronics)

$$R_s = \sqrt{R_c^2 + R_i^2}$$

So we have to find the intrinsic resolution to find the resolution of the system.

### 1.6.3 Experimental resolution

Another, equivalent way to find the resolution of the machine is to scan a little sphere filled of a certain quantity of tracer (a point source). The Full Width at Half Maximum (briefly FWHM) is the distance between the points

---

<sup>6</sup>it will be introduced in the next chapter.

<sup>7</sup>See [17] Ch 13, p.332 for a more detailed explanation.

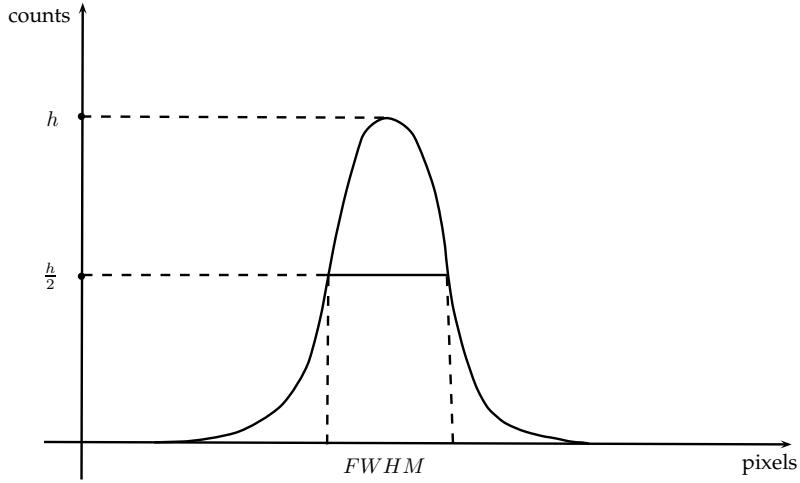


Figure 1.7: The FWHM of reconstructed data from a sphere filled with tracer. Note that is similar to a Gaussian function.

of level  $\frac{h}{2}$ , where  $h$  is the maximum of the function. E.g. if the function is the Gaussian  $f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right]$  it is easy to see that its FWHM is  $2\sqrt{2\log(2)}\sigma \approx 2.355\sigma$ .

The FWHM of the reconstructed data<sup>8</sup> offers a realistic, direct measure of the resolution of the machine. In fact two point sources distant less than FWHM are recognized as a unique, bigger source. On the other hand if their distance is greater than FWHM we can see the contribution of two different sources see Figure 1.8.

The resolution has also a role in partial volume effect: if the resolution is smaller, partial volume effect are less intense, because the sampling with a machine of a certain FWHM it is equivalent to the windowing with Gaussian function of  $\sigma = \frac{FWHM}{2\sqrt{2\log(2)}}$ : the windowed function will be smoother if  $\sigma$  is bigger.

#### 1.6.4 Sensitivity

The sensitivity of the detectors is given by the ratio between gamma ray passed through the collimators and all the gamma rays emitted by the source, i.e. is proportional to the solid angle subtended by irradiated detector area

---

<sup>8</sup>as we will see in the last chapter, the reconstruction method and the filtering of the image can change our estimate of the FWHM. In particular a low-pass filter increases the partial volume effect and the resolution rises by consequence.

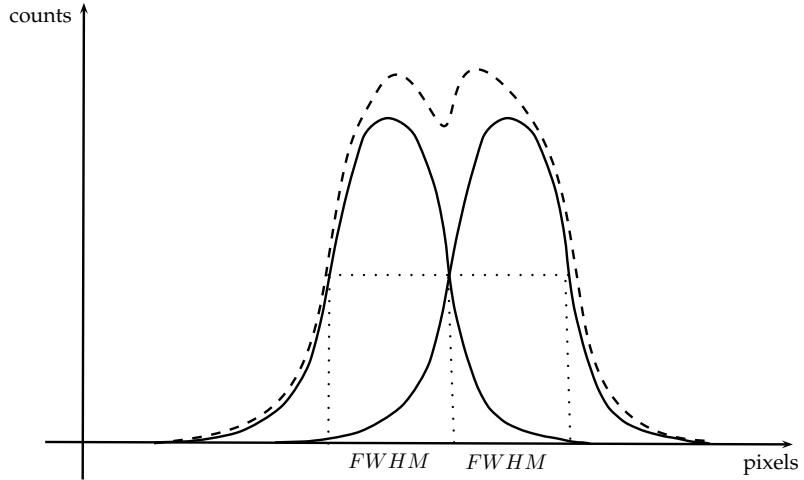


Figure 1.8: The FWHM is the limit for two objects to be recognizable as two; the sum of the two signals (dashed) at a distance  $R$  can be recognized as the contribution of two different objects.

multiplied by the square of the fraction of a collimator covered by the hole:

$$S \propto \frac{\pi R^2}{4\pi(L_{eff} + x)^2} \frac{D^2}{(D + t)^2},$$

$$S = K^2 \left( \frac{D^2}{L_{eff}(D + t)} \right)^2,$$

where  $K$  is a constant given by the shape of the collimator section ( $K = 0.28$  for square holes, 0.26 for hexagonal, 0.24 for round). Since  $D \gg t$  the formula becomes

$$S \cong \left( \frac{KD}{L_{eff}} \right)^2.$$

All these formulas will be used during the experiments.

## 1.7 Tracers used in SPECT and standard exams

The tracers used in SPECT are depending on the body part under examination and of the type of analysis to be performed.

Table 1.2 shows the main types of studies that use SPECT, and reports the respective tracer with its half-life, and some data about the projections and the size of the image that we want to get.

Type of Study	Substance	Emission Energy	Half-Life	#Angles	Size Image	Projection Time
Bone Scan	$^{99m}Tc$	140KeV	6h	120	128x128	30s
Myocardial perfusion scan	$^{99m}Tc$	140KeV	6h	60	64x64	25s
Brain scan	$^{99m}Tc$	140KeV	6h	64	128x128	30s
Neuroendocrine or neurological tumor scan	$^{123}I/^{131}I$	159KeV	13h/8days	60	64x64	30s

Table 1.2: Main types of studies that use SPECT tomography.

From this table we see that the size of the image to be reconstructed and the number of angles of sampling varies, but the typical values are 128x128 for the image size (or 64x64) and 120 (or 60) for the number of projections. Furthermore the time of acquisition for each angle is about 30s, hence the maximum exposure time is 60 minutes (that is also the limit of average patient's tolerance to stay inside the machine), which is far less than the half-life of tracers. For this we can assume that during the acquisition energy emitted by the tracer remains constant, i.e. does not depend on an acquisition time when it is conducted.

Let us take a look to the tracers. The  $^{99m}Tc$  technetium-99m is a nuclear isomer of the  $^{99}Tc$  ("m" stands for "metastable"); its half life is 6 hours, so the 93% of the atoms decay in 24 hours; moreover its radioactive emission is made up almost exclusively of gamma radiation with a single peak of energy (140KeV), sufficient for a good penetration through the body structures, but not dangerous for the patient. The  $^{99m}Tc$  has the ability to bind chemically in many biologically active molecules, e.g. when is combined with tin it binds blood cells and can therefore be used to discover eventual disorders of the circulatory system.

$^{123}I$  and  $^{131}I$  are radioisotopes of Iodine; they are used in radiotherapy, to cure hyperthyroidism and some types of thyroid cancer that absorb iodine. They emit gamma and beta rays so they can be seen in diagnostic scans after its use as therapy.

## 1.8 Some assumptions

In order to model correctly the tomography machine we will make some assumptions:

1. rays (X, gamma, etc..) All have no width, i.e. their length is much greater than their width, so we can consider them as straight lines.
2. incoming rays and detected rays are parallel to the line of angle  $\theta$  (parallel-beam); other cases must be specified.
3. the machine runs  $l$  scans at the angles  $(\theta_1, \dots, \theta_l)$ ; the detectors move on a circle (the angles are typically a uniform partition of  $[0, 2\pi]$  or  $[0, \pi]$  with a step of  $3^\circ$  or  $6^\circ$ ).
4. if the total time of acquisition is relatively short is indifferent to run an acquisition before or after another one<sup>9</sup>.
5. the reconstructed image has dimension  $N \times N$  (typically  $N = 128$  or  $N = 64$ ).
6. for each scan  $j \in 1, \dots, l$  the machine gets  $p$  linear samples  $(s_{1,j}, \dots, s_{p,j})$  (typically  $p = N$ , or  $p = 185$  or  $p = 95$ ). So the **sinogram** is an  $l \times p$  matrix of data collected in  $l$  scans each of  $p$  elements.
7. in the case of transmission tomography the ray which passes through the body is subject to an unknown attenuation  $f(x, y)$ , of which we know the sum along some lines at different angles.
8. in the case of emission tomography the rays start from the inside of the body and is subject to attenuation  $a(x, y)$  supposed to be note. Our goal is locating the radioactive concentration  $f(x, y)$ , given the attenuated sum along some lines at different angles.

---

<sup>9</sup>see previous section.



## Chapter 2

# Mathematical modeling: the Radon transform and its variants

Nella vita del signor Palomar c'è stata un'epoca in cui la sua regola era questa: primo, costruire nella sua mente un modello, il più perfetto, logico, geometrico possibile; secondo, verificare se il modello s'adatta ai casi pratici osservabili nell'esperienza; terzo, apportare le correzioni necessarie perché modello e realtà coincidano.

I. Calvino, *Palomar*



In this chapter different mathematical models of the tomography problem are given as transforms of functions, each coming from different assumptions on the projection method and on the kind of tomography performed.

### 2.1 The Beer-Lambert's law and the Radon transform

The **Radon transform**, introduced in 1917 by the Austrian mathematician Johann Karl August Radon, can be physically interpreted as a ray passing through different materials, each one of them with its own attenuation coefficient. We know the intensity in which the rays entered in and released by the body, so we can find the attenuation coefficients. In fact, Beer-Lambert's law says that if  $I(x)$  is the intensity of the ray,  $A(x)$  the attenuation coefficient at the point  $x$ , satisfies

$$\Delta I = -A(x)I(x) \Delta x.$$

Equivalently, by integrating their respective differential terms,

$$\int_{x_0}^{x_1} A(x)dx = - \int_{x_0}^{x_1} \frac{dI}{I} = - \log \left( \frac{I(x_1)}{I(x_0)} \right) \quad (2.1)$$

where the second term is known. Equation (2.1) has an equivalent on the 2-dimensional plane. In fact, given the change of coordinates due to rotation

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} t \\ s \end{pmatrix}$$

we get the formulation of the Radon transform as an integral on a line

$$\mathcal{R}f(t, \theta) = \int_{\ell_{(t, \theta)}} f = \int_{\mathbb{R}^2} f(\bar{x}) \delta(t - \bar{x} \cdot \bar{\theta}) d\bar{x}$$

where  $\bar{x} = (x, y)$ ,  $\bar{\theta} = (\cos \theta, \sin \theta)$ , or equivalently:

$$\mathcal{R}f(t, \theta) = \int_{\mathbb{R}} f(t \cos \theta - s \sin \theta, t \sin \theta + s \cos \theta) ds = \int_{\mathbb{R}} f(t\bar{\theta} + s\bar{\theta}^\perp) ds \quad (2.2)$$

where  $\bar{\theta}^\perp = (-\sin \theta, \cos \theta)$ .

## 2.2 Alternative Transforms

### 2.2.1 Backprojection

Now we may ask for every point  $(x, y)$  what is the average of the rays passing through that point. To answer this question we introduce the formal adjoint operator to  $\mathcal{R}$ , the so-called **backprojection operator**

$$\mathcal{R}^*g(x, y) = \frac{1}{|S^1|} \int_{S^1} g(x \cos \theta + y \sin \theta, \theta) d\theta = \frac{1}{|S^1|} \int_{\mathbb{R} \times S^1} g(s, \theta) \delta(s - \bar{x} \cdot \bar{\theta}) ds d\theta \quad (2.3)$$

or equivalently

$$\mathcal{R}^*g(x, y) = \frac{1}{|S^1|} \int_{S^1} g(\bar{x} \cdot \bar{\theta}, \theta) d\theta$$

where  $S^1$  is the unit hemisphere represented by the parameter  $\theta$ , i.e.  $S^1 = [0, \pi]$  or  $S^1 = [0, 2\pi]$  according to the angular sampling, so  $|S^1| = \pi$  or  $2\pi$ .

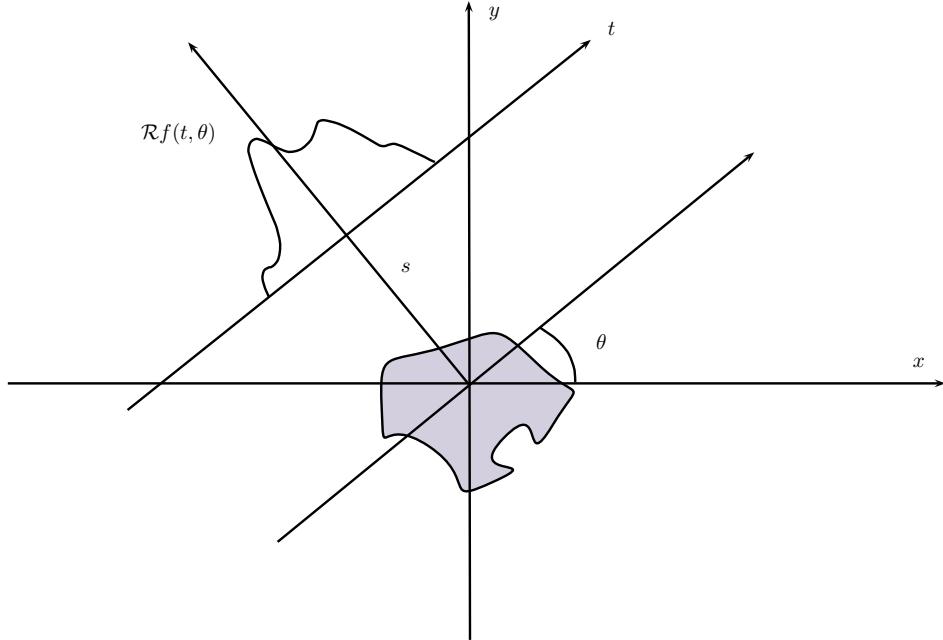


Figure 2.1: Radon transform of an object with constant attenuation coefficient.

### 2.2.2 Divergent beam transform

In the Radon transform formulation we assumed that the collimators receive only the parallel rays. Rays are not parallel, but start from a single transmitter and are disposed radially. If now we let pass through the collimators all the rays, we have to consider the **fan beam transform**, also called **divergent beam transform**

$$\mathcal{D}h(\bar{x}, \theta) = \int_0^{+\infty} h(x + t \cos \theta, y + t \sin \theta) dt = \int_0^{+\infty} h(\bar{x} + t\bar{\theta}) dt \quad (2.4)$$

where the parameter  $t$  this time indicates the distance traveled by the ray from the emitter.

### 2.2.3 Attenuated Radon transform

In the case of an emission tomography, as PET or SPECT, modeling the tomography becomes more difficult: in fact each single beam emitted by the tracer in the highlighted area must pass through objects of different consistency, and then will be subject to attenuation. From Beer-Lambert's law we obtain that

$$I(x_1) = I(x_0) \exp \left( - \int_{x_0}^{x_1} A(x) dx \right).$$

We assume then to know the attenuation coefficients<sup>1</sup>  $a(x, y)$ , we know that a ray passing through  $(x, y)$  changes direction with a certain probability directly proportional to  $a(x, y)$ . Then, if  $f(x, y)$  is the concentration of the tracer, then the radiation detected at an angle  $\theta$  to a first approximation is given by

$$\mathcal{R}^a f(t, \theta) = \int_{\ell_{(t, \theta)}} e^{-a(\bar{x})} f(\bar{x}) = \mathcal{R}(e^{-a} f)(t, \theta). \quad (2.5)$$

which is simply the Radon transform of a (weighted) function  $e^{-a} f$ .

We can obtain a more realistic version of (2.5) if we consider that a ray subjected to diffraction expands radially from the point of diffraction

$$\mathcal{R}_a f(t, \theta) = \int_{\ell_{(t, \theta)}} e^{-\mathcal{D}a(\bar{x}, \theta + \frac{\pi}{2})} f(\bar{x}) \quad (2.6)$$

This operator is called **attenuated Radon transform** (briefly AtRT), where  $\mathcal{D}$  is the divergent beam transform.

Note that this new transform is **not** equivalent to the Radon transform of a modified of weighted function.

In the special case of PET tomography, where the positrons come out two by two in opposite directions, the attenuation part in (2.6) goes outside the integral<sup>2</sup> and then to recover the section we only have to apply an **attenuation correction** to the image obtained by the inversion of the classic Radon transform to obtain a realistic reconstruction. If the scan is done on a relatively homogeneous areas such as the brain, we can also consider the attenuation as a constant<sup>3</sup>.

---

<sup>1</sup>the attenuation coefficients were previously unknown and called  $f(x, y)$ , now we call them  $a(x, y)$  and we suppose to know them, while the unknown function is now  $f(x, y)$ , the tracer's concentration

<sup>2</sup>see [13] Ch 1.2, p.4, and [18] Ch 6, p.91 for more details.

<sup>3</sup>more details on the dedicated chapter.

## Chapter 3

# Analytical properties and inversion of the Radon transform

[...] as if all this had been done and said so many times before it made you feel it was recorded, they all in here existed basically as Fourier Transforms of postures and little routines, locked down and stored and call-uppable for rebroadcast at specified times.

D.F. Wallace, *Infinite Jest*



et us show some properties of the Radon transform and its related transforms; some of this properties will be useful in the following chapters, some other will lead us to an analytical resolution formula to the following problem

$$\text{Given } g \text{ find } f \text{ such that } \mathcal{R}f = g \quad (3.1)$$

Let us begin with some definitions:

**Definition 1.** A function  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$  is called **(Radon-)transformable** if  $\mathcal{R}f(t, \theta)$  exists and is finite  $\forall (t, \theta) \in \mathbb{R} \times S^1$ .

We will not go deeper in discussing the domain of the Radon transform, but it's good to know from [9] and [13] that the Radon transform can be seen both as a map from the Schwartz space  $\mathcal{S}$  in a subset of  $\mathcal{S}$ , or from  $L^2$  to itself or from  $C_0^\infty$  to itself<sup>1</sup>. Let us define in general

**Definition 2.** A function  $f$  is **transformable** (as regards one of the operators defined in previous chapter) if its transform exists and is finite (almost) everywhere.

---

<sup>1</sup>most of theorems shown below require  $f \in \mathcal{S}$  or  $f \in C_0^\infty$ .

### 3.1 Basic properties

**Lemma 3.1.1.** *Let  $f$  be a Radon-transformable function. Hence*

$$\mathcal{R}f(t, \theta) = \mathcal{R}f(-t, \pi + \theta)$$

*Proof.* Trivial by substitution in (2.2) of page 16.  $\square$

We can therefore restrict the range of angular projections in  $\theta \in [0, \pi]$ , as the projections from the angles from  $\pi$  to  $2\pi$  do not provide new data; on the contrary in the case of attenuated transform the sampling must be made of angles from 0 to  $2\pi$ .

**Lemma 3.1.2.** *The operators  $\mathcal{R}$ ,  $\mathcal{R}^*$ ,  $\mathcal{D}$ ,  $\mathcal{R}_a$  are linear on their respective spaces of transformable functions.*

*Proof.* This is also trivial by the linearity of the integral operator. E.g.

$$\begin{aligned} \mathcal{R}(\alpha f + \beta g)(t, \theta) &= \int_{\mathbb{R}^2} (\alpha f(\bar{x}) + \beta g(\bar{x})) \delta(t - \bar{x} \cdot \bar{\theta}) d\bar{x} = \\ &= \alpha \int_{\mathbb{R}^2} f(\bar{x}) \delta(t - \bar{x} \cdot \bar{\theta}) d\bar{x} + \beta \int_{\mathbb{R}^2} g(\bar{x}) \delta(t - \bar{x} \cdot \bar{\theta}) d\bar{x} = \alpha \mathcal{R}f(t, \theta) + \beta \mathcal{R}g(t, \theta) \end{aligned}$$

$\square$

**Proposition 3.1.3.** *For every appropriated functions  $f$  and  $g$ , then*

$$\mathcal{R}(f *_2 g) = \mathcal{R}f * \mathcal{R}g$$

$$\mathcal{R}^*(g * \mathcal{R}f) = \mathcal{R}^*g * \mathcal{R}f$$

where “\*” is meant to be the convolution made only on the linear variable  $t$  and “\*\_2” the 2-dimensional convolution.

*Proof.* see [13] Ch 2.1, p.13-14 for the proof.  $\square$

Next Lemmas will be useful in interpolation.

**Lemma 3.1.4.** *If  $f$  is transformable and radial, i.e  $f(\bar{x}) = \varphi(\|\bar{x}\|)$ , then  $\mathcal{R}f$  only depends on the linear parameter  $t$*

*Proof.*

$$\begin{aligned} \mathcal{R}f(t, \theta) &= \int_{\mathbb{R}} f(t\bar{\theta} + s\bar{\theta}^\perp) ds = \int_{\mathbb{R}} \varphi(\|t\bar{\theta} + s\bar{\theta}^\perp\|) ds = \\ &= \int_{\mathbb{R}} \varphi \left( \left\| \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} t \\ s \end{pmatrix} \right\| \right) ds = \int_{\mathbb{R}} \varphi(\sqrt{t^2 + s^2}) ds \end{aligned}$$

because the matrix is orthogonal.  $\square$

**Theorem 3.1.5. (Shift property)** Let  $\bar{c}$  be a constant vector and  $g(\bar{x}) = f(\bar{x} - \bar{c})$ . Then

$$\mathcal{R}g(t, \theta) = \mathcal{R}f(t - \bar{c} \cdot \bar{\theta}, \theta)$$

*Proof.* See [16] Ch 2.7, p.31.  $\square$

This last Theorem guarantees the uniqueness of the reconstruction problem (3.1) in the continuous case.

**Theorem 3.1.6. (Uniqueness)** Let  $f \in \mathcal{S}$  and  $K \subset \mathbb{R}^2$  be a convex, compact set called **the hole**. If  $\mathcal{R}f = 0$  in every plane  $\{\bar{x} \cdot \bar{\theta} = s\} \cap K = \emptyset$ . Then  $f = 0$  outside  $K$ .

*Proof.* Proof can be found in [13] Ch 2.3, p.30-32.  $\square$

## 3.2 Sampling and resolution

In order to solve the problem of reconstruction (3.1) in the discrete case we have to make some further assumptions on the function  $f$ . We shall also make some assumptions on the sampling to avoid aliasing problems. Let us begin with a definition and a standard Theorem in applied mathematics

**Definition 3.** A Fourier-transformable function  $f$  is called  **$b$ -band-limited** if its Fourier transform  $\mathcal{F}f(\nu)$  is negligible for  $|\nu| > b$ , i.e.

$$\int_{|\nu|>b} |\mathcal{F}f(\sigma)| d\sigma = 0$$

**Theorem 3.2.1. (Shannon-Nyquist)** Let  $f$  be a  $b$ -band-limited function and let  $h \leq \pi/b$ . Then  $f$  is uniquely determined by the values  $f(hk)$ ,  $k \in \mathbb{Z}^n$  and in  $L^2(\mathbb{R}^n)$

$$f(x) = \sum_k f(hk) \text{sinc}\left(\frac{\pi}{h}(\bar{x} - hk)\right)$$

moreover we have

$$\|f\|_{L^2} = h^{n/2} \left( \sum_k |f(hk)|^2 \right)^{1/2}$$

*Proof.* the proof is shown in [13] Ch 3.1, p.56.  $\square$

This theorem gives a condition to recover efficiently a  $b$ -band-limited function: the sampling step  $h$  must be smaller than  $\pi/b$  (this is called **Nyquist condition**).

Now we want to find a condition on the angular step for recover in a reliable way the function  $f$ . Let us proceed with some further definitions and properties.

**Definition 4.** A *spherical harmonic function* of degree  $m$  is the restriction to  $S^1$  of harmonic, homogeneous polynomial function of degree  $m$  on  $\mathbb{R}^2$

**Proposition 3.2.2.** If  $H_m$  is the space of the spherical harmonic functions of degree  $\leq m$  in  $\mathbb{R}^2$  then a basis for this space is

$$\{p_0, p_{(l,1)}, p_{(l,-1)}\}_{l=1,\dots,m}$$

with

$$\begin{aligned} p_0 &= 1 \\ p_{(l,1)}(\bar{\theta}) &= \cos(l\theta) \\ p_{(l,-1)}(\bar{\theta}) &= \sin(l\theta) \end{aligned}$$

where, as usual  $\bar{\theta} = (\cos \theta, \sin \theta)$

We can easily verify that  $p_{(l,\pm 1)}$  are harmonic and homogeneous polynomials in  $\bar{\theta}$ ; in fact, by de Moivre's formula

$$\cos(nx) + i \sin(nx) = (\cos x + i \sin x)^n$$

we can easily find the explicit functions by taking

$$p_{(l,\pm 1)} = \Re(\cos \theta + i \sin \theta)^l$$

developed with the Newton's binomial formula. Then these functions are all homogeneous polynomials in  $(\cos \theta, \sin \theta) = (x, y)$ . On the other side these functions are harmonic too, since their laplacians are

$$\begin{aligned} \Delta(p_{(l,+1)}) + i\Delta(p_{(l,-1)}) &= \Delta(p_{(l,+1)} + ip_{(l,-1)}) = \Delta((x + iy)^n) = \\ &= n(n-1)(x + iy)^{n-2} + (i)^2 n(n-1)(x + iy)^{n-2} = 0 \end{aligned}$$

Now we define a subset of this space which will be useful for next theorems

**Definition 5.** The space  $H'_m$  is the subset of the functions in  $H_m$  which are even function for  $m$  even and odd functions for  $m$  odd.

We observe immediately that  $\dim(H'_m) = m + 1$  because the basis of this space is  $\{p_{(l,(-1)^l)}\}_{l=0\dots m}$ .

**Definition 6.** *A  $\subset S^1$  is ***m*-resolving** if the only function on  $H'_m$  vanishing in  $A$  is the function  $p = 0$ .*

**Lemma 3.2.3.** *If  $A$  is *m*-resolving then  $|A| \geq \dim(H'_m) = m + 1$ .*

**Theorem 3.2.4.** *Let  $A$  be a *m*-resolving set and  $f \in C_0^\infty(B(0,1))$ . If  $\mathcal{R}f$  vanishes in  $A$ , then for  $0 < \vartheta < 1$*

$$\|f\|_{L^\infty} \leq \frac{1}{2\pi\eta(\vartheta, m)} \varepsilon_0(f, \vartheta m)$$

where  $\eta$  is a function such that

$$0 < \eta(\vartheta, m) < C(\vartheta) \exp(-\lambda(\vartheta)m)$$

and  $b \geq B(\vartheta)$ ,  $B, C, \lambda$  positive, and

$$\varepsilon_0(f, \vartheta m) = \frac{1}{2\pi} \int_{|\sigma| \geq \vartheta m} |\mathcal{F}g(\sigma, \theta)| d\sigma$$

*Proof.* see [13] Ch 3.3, p.68 for the proof and for further details.  $\square$

This last theorem gives an answer to the question of resolution. In fact if  $f$  is  $b$ -band-limited then  $\varepsilon_0$  is negligible and  $f$  can be recovered reliably from a  $m$ -resolving set, where  $m > b$ , i.e. where the angular sampling points are at least  $l = m + 1 = \lfloor b \rfloor + 2$ .

**Definition 7.** *Let  $f \in C_0^\infty(B(0,1))$  be  $b$ -band limited and such that  $\mathcal{R}f$  vanishes on a  $l$ -resolving set. A sampling of  $\mathcal{R}f$  of the form  $\{(t_i, \theta_j)\}$  of linear step  $h$  and  $l$  angular data, is said that  $\mathcal{R}f$  is ***reliably sampled*** if  $l - 1 > b$  and  $h < \pi/b$ .*

In conclusion if we take  $l$  angular samples of the projection data each one made by  $p$  linear equispaced acquisitions of step  $h = \frac{2}{\lfloor p \rfloor}$ , then the maximum frequency of the data will be

$$b = \min\{l - 2, \frac{\pi}{2}p\} \tag{3.2}$$

and the resolution will be

$$R_{sampling} = \frac{2\pi}{b}$$

with attention to report the radius of sampling to  $r$  instead of 1, i.e. we have to multiply  $R_{sampling}$  by  $r$ .

### 3.3 Inversion formulas for the Radon transform

**Proposition 3.3.1.** *The backprojection operator is the Hermitian added to the transform  $\mathcal{R}$ , but not its inverse: in fact*

$$\mathcal{R}^* \mathcal{R} f(\bar{x}) = 2 \int_{\mathbb{R}^2} \frac{f(\bar{y})}{|\bar{x} - \bar{y}|} d\bar{y} = \frac{2}{|\bar{x}|} *_2 f$$

*Proof.* see [13] Ch 2.1, p.15 for the proof.  $\square$

this means that what we get from applying the backprojection operator to projection data is a *blurred* image. From this property we can derive a first numerical method of inversion which simply uses the 2D FFT and IFFT algorithms for a fast deconvolution. The computational complexity of such algorithm is known to be  $O(N^2 \log N)$  where  $N$  is the number of pixel by size in the reconstructed image.

**Theorem 3.3.2. (Central Slice Theorem)** *Let  $f$  be a Fourier and Radon transformable function. Then*

$$\mathcal{F}_2 f(T \cos \theta, T \sin \theta) = \mathcal{F}(\mathcal{R} f)(T, \theta)$$

Where we mean for  $\mathcal{F}_2$  the 2-dimensional Fourier transform and  $\mathcal{F}$  the 1-dimensional Fourier transform applied only to the variable  $t$ .

*Proof.* proof can be found in [5] Ch 6, p.47; a more general result can be found in [13] Ch 2.1, p.10.  $\square$

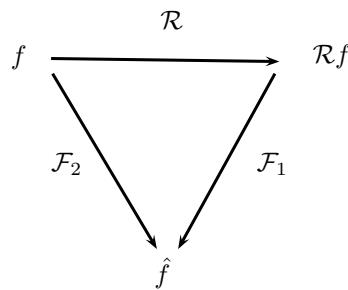


Figure 3.1: Scheme for Central Slice Theorem: the 2-dimensional Fourier transform of a function in polar coordinates is equal to the 1-dimensional Fourier transform made on the linear parameter of its Radon transform.

**Theorem 3.3.3. (Inverse Radon transform)** *The original function can be derived analytically using the **filtered backprojection** formula*

$$f = \frac{1}{2} \mathcal{R}^* [\mathcal{F}^{-1}(|\nu| \mathcal{F}(\mathcal{R}f))] \quad (3.3)$$

where we mean that the direct and inverse Fourier transform is applied only to the variable  $t$ .

*Proof.* see [5] Ch 6, p.49 or [18] Ch 7, p.46 for the proof.  $\square$

In what follows we will refer this formula as FBP.

Given the inaccurate numerical results of this formula, we use

$$w(\nu) = v(\nu)|\nu|$$

instead of  $|\nu|$ , with  $v$  a **low-pass filter**<sup>2</sup>. The reconstruction formula becomes

$$f \cong \frac{1}{2} \mathcal{R}^* [\mathcal{F}^{-1}(w(\nu) \mathcal{F}(\mathcal{R}f))] \quad (3.4)$$

Sometimes we know  $\tilde{w}(t) = \mathcal{F}^{-1}(w(\nu))$  in order to apply the convolution directly to the  $\mathcal{R}f$ , without going through the Fourier transforms, i.e.

$$f \cong \frac{1}{2} \mathcal{R}^* [\tilde{w}(t) * \mathcal{R}f]$$

An example of filter is the function  $w(\nu) = \chi_{[-L,L]}|\nu|$  with  $L$  fixed.

### 3.4 Discretization of the FBP formula and error bound

We can get another approach proposed by [13] in Ch 5.1 for the resolution of the problem (3.1). In fact thanks to Theorem (3.1.3), if  $W_b = \mathcal{R}^*(w_b)$ , then

$$W_b * f = \mathcal{R}^*(w_b * \mathcal{R}f)$$

if we assume that  $w_b$  is such that  $W_b$  approximates the  $\delta$ -distribution we get

$$f \cong \mathcal{R}^*(w_b * \mathcal{R}f) \quad (3.5)$$

which is equivalent to (3.4). As said in the previous section we will choose  $W_b$  as a low-pass filter with  $b$  its cut-off frequency.

---

<sup>2</sup>for a list of the low-pass filter used see the chapter about regularization.

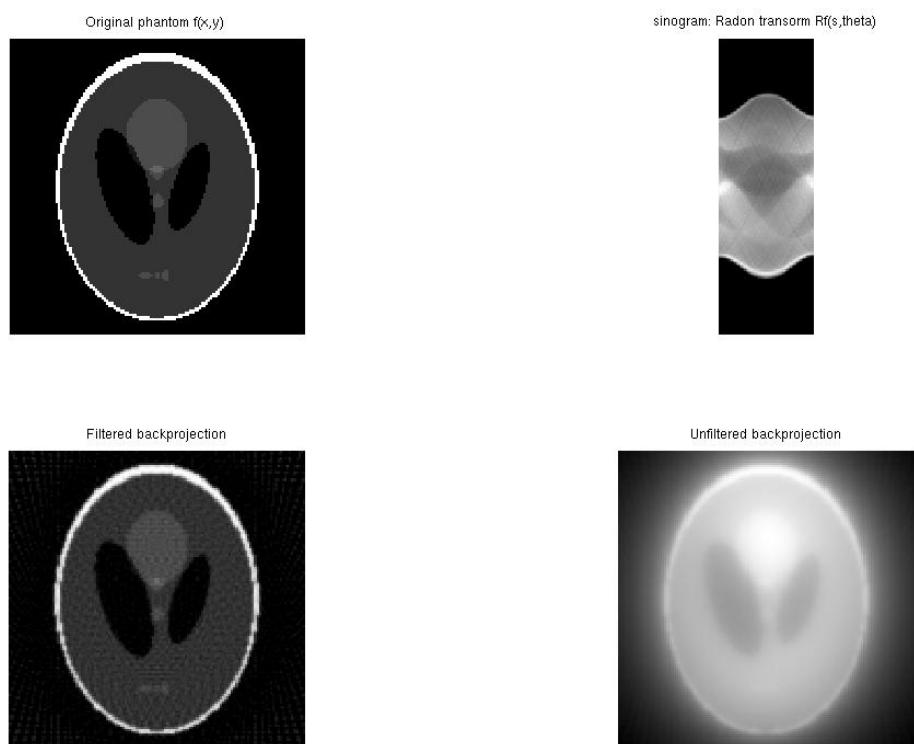


Figure 3.2: Original Shepp-Logan phantom, sinogram of its Radon transform, and reconstruction by filtered and unfiltered back projection formulas.

Now we consider the discrete version of (3.1). Let  $g = \mathcal{R}f$  be sampled at  $(t_i, \theta_j)$  with  $i = -q, \dots, q$  and  $j = 1, \dots, l$  and let  $h = 1/q$  be the linear step, i.e.  $t_i = hi$ ; in this notation  $p = 2q + 1$ . If we consider the discrete convolution

$$a \stackrel{h}{*} b(t) = h \sum_{i=-q}^q a(t - t_i) b(t_i)$$

and the discrete backprojection given from a quadrature rule

$$\mathcal{R}_p^* g(\bar{x}) = \frac{1}{\pi} \sum_{j=1}^p \alpha_{(p,j)} g(\bar{x} \cdot \theta_j, \theta_j)$$

with  $\alpha_{(p,j)}$  positive weights, now the formula (3.5) reads as

$$f_{FBP} = \mathcal{R}_p^* (w_b \stackrel{h}{*} \mathcal{R}f) \quad (3.6)$$

The following theorem gives an error estimate for this last formula.

**Theorem 3.4.1.** *Let  $f \in C_0^\infty(B(0, 1))$ , assume that the quadrature rule used before is exact in  $H'_{2m}$ , and let  $\mathcal{R}f$  be reliably sampled, i.e., for some  $\vartheta \in (0, 1)$*

$$b < \vartheta m \quad b < \pi/h$$

then

$$f_{FBP} = W_b * f + e_1 + e_2$$

and

$$|e_1| \leq \frac{1}{4\pi} \varepsilon^*(f, b)$$

$$|e_2| \leq \eta(\vartheta, m) \|f\|_{L^\infty}$$

where

$$\varepsilon^*(f, b) = |S^1| \sup_{\theta \in S^1} \int_{|\sigma| \geq b} |\mathcal{F}g(\sigma, \theta)| d\sigma$$

and  $\eta$  is any function such that

$$0 < \eta(\vartheta, m) < C(\vartheta) \exp(-\lambda(\vartheta)m)$$

and  $b \geq B(\vartheta)$ ,  $B, C, \lambda$  positive.

*Proof.* see [13] Ch 5.1, p.104 for the proof.  $\square$

Let us try to explain the meaning of this theorem. If  $f$  is  $b$ -band-limited, the term  $e_1$  is negligible, so the first error, related to the inaccuracy of the discretization of the convolution product, is due to aliasing problems. On the other side the second error term  $e_2$  depends on the chosen discretization and the discretization of the backprojection operator.

Moreover from the proof<sup>3</sup> of Theorem (3.4.1) we can find a weaker result which will be more useful for our purposes

$$|e_2| \leq 2|S^1| \|w_b * g - (w_b * g)_m\|_{L^\infty}$$

where  $g = \mathcal{R}f$  and  $(w_b * g)_m$  is the truncation to order  $m = b$  of the spherical harmonics series. Since  $w_b$  is given exactly we can assume

$$w_b * g - (w_b * g)_m = w_b * (g - \tilde{g})$$

with  $\tilde{g}$  the truncation to order  $b$  of the spherical harmonics series for  $g$ , i.e.  $g - \tilde{g}$  is the error on the approximation of  $g = \mathcal{R}f$ .

**Lemma 3.4.2.** *For every suitable functions  $a$  and  $b$  on a suitable set  $\Omega \subseteq \mathbb{R}^s$ ,  $s \in \mathbb{N}$ , we get*

$$\|a * b\|_{L^\infty} \leq \|a\|_{L^1} \|b\|_{L^\infty}$$

*Proof.*

$$\begin{aligned} \|a * b\|_{L^\infty} &= \sup_x \left| \int_{\Omega} a(x - y) b(y) dy \right| \leq \sup_x \int_{\Omega} |a(x - y)| |b(y)| dy \leq \\ &\leq \sup_x |b(x)| \sup_x \int_{\Omega} |a(x - y)| dy = \sup_x |b(x)| \int_{\Omega} |a(y)| dy = \|a\|_{L^1} \|b\|_{L^\infty} \end{aligned}$$

□

**Remark:** this is a particular case of Young's inequality for convolutions.

We are now able to prove the following Theorem

**Theorem 3.4.3.** *Let  $f \in C_0^\infty(B(0, 1))$  a  $b$ -band-limited function, and let  $g = \mathcal{R}f$  be reliably sampled, i.e.*

$$b < m \quad \text{and} \quad b < \pi/h$$

*let  $\tilde{f}$  be the FBP reconstruction, then*

$$|f(\bar{x}) - \tilde{f}(\bar{x})| \leq 2|S^1| \|w_b\|_{L^1} \|g - \tilde{g}\|_{L^\infty} + |e_3|$$

---

<sup>3</sup>this proof is available in [13], as said previously.

$\forall \bar{x} \in B(0, 1)$ , i.e.

$$\|f - \tilde{f}\|_{L^\infty(\mathbb{R}^2)} \leq 2|S^1| \|w_b\|_{L^1(\mathbb{R})} \|g - \tilde{g}\|_{L^\infty(\mathbb{R} \times S^1)} + |e_3|$$

with  $e_3$  the quadrature error which depends on the interpolation chosen to approximate the backprojection integral.

We will see some application of this Theorem in the Chapter of the experiments.



## Chapter 4

# Algebraic Reconstruction Techniques

Insomma il secondo t0 in cui stanno la freccia F0 e un po' più in là il leone L0 e qui il me stesso Q0 è uno strato spaziotemporale che resta fermo e identico per sempre, e accanto ad esso si dispone t1 con la freccia F1 e il leone L1 e il me stesso Q1 che hanno leggermente cambiato le loro posizioni, e lì affiancato c'è t2 che contiene F2 e L2 e Q2 e così via. In uno di questi secondi messi in fila risulta chiaro chi vive e chi muore tra il leone Ln, e il me stesso Qn, e nei secondi seguenti stanno certamente svolgendosi [...]. Ogni secondo è definitivo, chiuso, senza interferenze con gli altri.

I. Calvino, *Ti con zero*

 e will now introduce a completely different approach to the tomography problem. Let us consider the image reconstruction as a problem of interpolation. In fact, given a basis of functions  $\{b_i(\bar{x})\}_{i=1\dots n}$  that interpolate the function  $f$ , i.e. such that

$$f(\bar{x}) = \sum_{i=0}^n c_i b_i(\bar{x}) \quad \forall \bar{x} \in X$$

where  $X$  is properly chosen, then for linearity of the Radon transform shown in Lemma (3.1.2)

$$\mathcal{R}f(\bar{y}) = \sum_{i=0}^n c_i \mathcal{R}b_i(\bar{y}) \quad \forall \bar{y} \in Y = \{(t_j, \theta_j)\}.$$

Now consider the problem from a practical point of view: incoming data are stored in a matrix  $l \times p$  where  $l$  is the number of angles of acquisition and  $p$  the number of pixels recorded for each angle  $(\theta_1, \dots, \theta_l)$ . We want to obtain as output an image of  $N \times N$  pixels. Let us consider now the domain as this grid of pixels: let  $P_i = [x_i, x_{i+1}] \times [y_i, y_{i+1}]$  the  $i$ -th pixel, where  $(x_i, y_i)$  is its top-left vertex. Then we introduce the following basis

$$b_i(x, y) = \chi_{P_i}(x, y).$$

We know the Radon transform of each one of these items

$$\mathcal{R}b_i(t, \theta) = \text{meas}(\ell_{t, \theta} \cap P_i)$$

where  $\ell_{t, \theta} = \{t\bar{\theta} + s\bar{\theta}^\perp \mid s \in \mathbb{R}\}$ ; hence this integral can explicitly calculated for every angle chosen.

The obtained system is

$$Ax = c \tag{4.1}$$

where  $A(i, j) = \mathcal{R}b_i(t_j, \theta_j)$  is a sparse matrix  $N^2 \times lp$ ,  $x$  is the unknown vector such that  $x_i = f(x_i)$  and  $c$  is the vector of the projection data  $c_j = \mathcal{R}f(t_j, \theta_j)$ . The methods using this approach are known as **Algebraic Reconstruction Techniques** or **ART**.

## 4.1 Some iterative methods

Nowadays the obtained system is solved numerically in a large amount of time: in fact if the matrix has standard dimensions ( $N = 128$ ,  $l = 60$ ,  $p = 128$ ) and to store it we need more than 1Gb of memory. The problem can be bypassed considering the sparse structure of the matrix  $A$  (at least 90 % of its elements is equal to 0). For solving the system (4.1) we also will use iterative algorithms, whose general structure is the following:

- the initial vector  $x^{(0)}$  is a *blank* (the zero vector or unit vector) or a inaccurate reconstruction.
- the image at step  $k$ ,  $x^{(k)}$ , is projected and compared with the data (i.e., for instance,  $e^{(k)} = c - Ax^{(k)}$ ).
- the image is modified considering the error found in the previous step ( $x^{(k+1)} = g(x^{(k)}, e^{(k)})$  with  $g$  a certain function).

Here are some of the most common algorithms:

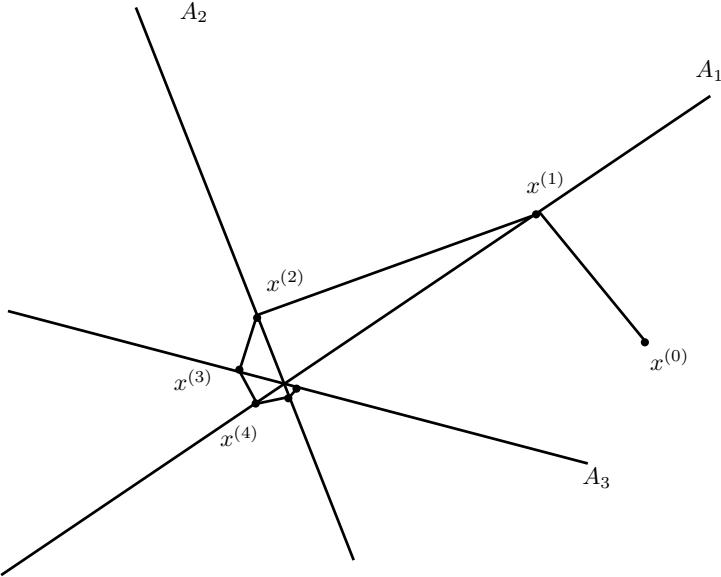


Figure 4.1: The main idea behind Kaczmarz algorithm is the projection on the columns of the matrix  $A$ ; in the figure we used three columns as example.

#### 4.1.1 The Kaczmarz algorithm

Let us choose iteratively a row  $k$  of the matrix  $A$  (usually  $i \equiv k \bmod lp$  or we choose randomly). Then we project the image on the vector  $A_k$ .

$$x^{(k+1)} = x^{(k)} + \frac{c_i - A_i^T \cdot x^{(k)}}{A_i^T \cdot A_i} A_i$$

In practice using this algorithm we reach a satisfying reconstruction only after having projected the image on all the rows of the matrix. In standard dimensions, the solution is accurate only after  $O(10^4)$  iteration. That is why, in order to increase the speed convergence, sometimes we adopt a relaxation parameter  $\lambda$ , chosen properly or randomly (or with probability proportional to  $\|A_i\|^2$ ) and the iterative scheme becomes

$$x^{(k+1)} = x^{(k)} + \lambda_k \frac{c_i - A_i^T \cdot x^{(k)}}{A_i^T \cdot A_i} A_i$$

in this case the method is called **Kaczmarz algorithm**. It is shown in [18], [13] Ch 5.3 p130-132 and [10] Ch 2 p.2 that the convergence is guaranteed for  $\lambda_k \in (0, 2)$ . The ART algorithm is the first method historically used, but it was abandoned for the FBP because of its slowness of convergence.

### 4.1.2 Maximum Likelihood Expectation Maximization

The main idea of this algorithm<sup>1</sup>, based on probabilistic arguments, is to find the  $x$  that maximizes the likelihood of the probability  $L(x) = P(c|x)$ . This means that we are looking for the solution  $x$  which maximizes the conditioned probability of having our data  $c$  with an image  $x$ .

Assuming that the noise is Poissonian on  $c$ , then the likelihood is

$$L(x) = P(c|x) = \prod_{i=1}^{lp} \frac{(c_i^*)^{c_i}}{c_i!} e^{-c_i^*}$$

where  $c^* = Ax$  is the exact sinogram, i.e. the projection of the exact solution  $x$ . Equivalently, it maximizes

$$l(x) = \log(L(x)) = \sum_{i=1}^{lp} -(Ax)_i + C_i \log(Ax)_i + K$$

with  $K$  a constant and constraint non-negativity on  $x_i$ . Kuhn-Tucker conditions are then:

$$\begin{aligned} x_j \frac{\partial l(x)}{\partial x_j} &= 0 \text{ if } x_j > 0 \\ \frac{\partial l(x)}{\partial x_j} &\leq 0 \text{ if } x_j = 0 \end{aligned}$$

so that we get from the first equation

$$0 = x_j \frac{\partial l(x)}{\partial x_j} = x_j \left( - \sum_{i'=1}^{lp} a_{i',j} + \sum_{i=1}^{lp} \frac{a_{i,j} c_i}{\sum_{j'=1}^{N^2} a_{i,j'} x_{j'}} \right)$$

or equivalently

$$x_j = \frac{x_j}{\sum_{i'=1}^{lp} a_{i',j}} \sum_{i=1}^{lp} \frac{a_{i,j} c_i}{\sum_{j'=1}^{N^2} a_{i,j'} x_{j'}}$$

from which comes out the iterative scheme

$$x_j^{(k+1)} = \frac{x_j^{(k)}}{\sum_{i'=1}^{lp} a_{i',j}} \sum_{i=1}^{lp} \frac{a_{i,j} c_i}{\sum_{j'=1}^{N^2} a_{i,j'} x_{j'}^{(k)}}$$

---

<sup>1</sup>see [18] for a more detailed explanation.

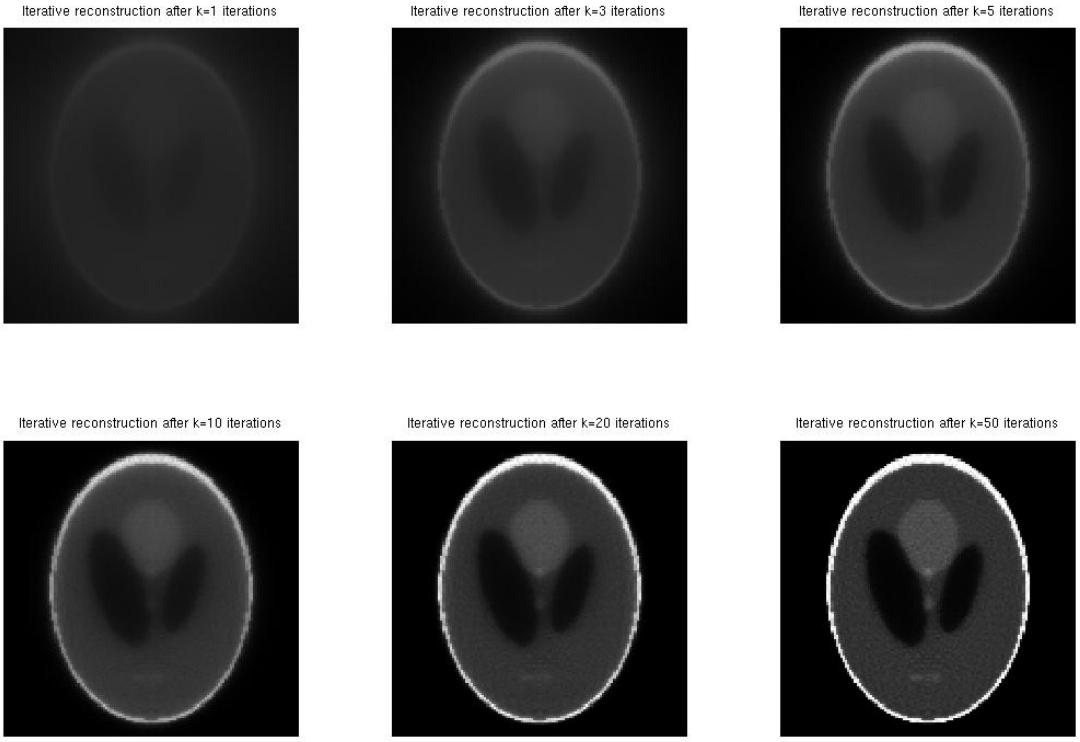


Figure 4.2: Several iterations in MLEM reconstruction of a phantom.

We now rewrite this scheme in a vectorial and multi-step form

1.  $c^f := Ax^{(k)}$
2.  $c^q := c./c^f$  (punctual division)
3.  $x^b = A^T c^q$
4.  $s_j = \sum_i a_{i,j}$
5.  $x^{(k+1)} = x^{(k)}. * x^b ./ s$  (product and division are made elementwise)

Faster results are obtained by using the **OSEM** algorithm (Ordered Subset Expectation Maximization); we simply apply EM reconstruction of ordered subset of all the projections  $\{\theta_i\}_{i=1\dots l} = \Theta = \bigcup_j \Theta_j$ .

**The OSEM algorithm is currently the most used in nuclear medicine for its speed of convergence and accuracy.**

### 4.1.3 Conjugate gradient

The conjugate gradient method is widely used in all fields for solving linear systems  $Ax = b$  with  $A$  square matrix of order  $n$ , symmetric and positive definite. In our framework the method is known as **LSCG (Least Squares Conjugate Gradient)** because we apply the conjugate gradient method to the normal equation of the system (4.1).

In fact the matrix  $A^T A$  is square, symmetric and positive definite. The method proposed below solves iteratively the normal equations

$$A^T A x = A^T c$$

without having to compute explicitly  $A^T A$ . The algorithm starts with an initialization phase:

- given  $x^{(0)}$  initial value
- $s^{(0)} = c - Ax^{(0)}$
- $r^{(0)} = p^{(0)} = A^T s^{(0)}$
- $q^{(0)} = Ap^{(0)}$

then begins to calculate, at each step:

- $\alpha = \frac{|r^{(k)}|^2}{|q^{(k)}|^2}$
- $x^{(k+1)} = x^{(k)} + \alpha p^{(k)}$
- $s^{(k+1)} = s^{(k)} - \alpha q^{(k)}$
- $r^{(k+1)} = A^T s^{(k)}$
- $\beta = \frac{|r^{(k+1)}|^2}{|r^{(k)}|^2}$
- $p^{(k+1)} = r^{(k+1)} + \beta p^{(k)}$
- $q^{(k+1)} = Ap^{(k+1)}$

This algorithm, despite the appearance, is in most cases as fast as an iteration of EM.

## 4.2 Accelerated reconstruction

Although iterative algorithms are extremely fast, it is possible to further accelerate the reconstruction, by using a technique of *zero-padding* of matrix  $A$ , i.e.

$$\tilde{a}_{(i,j)} = \begin{cases} a_{i,j} & \text{if } a_{i,j} > \gamma \\ 0 & \text{else} \end{cases}$$

The algorithm is trivial,

$$A = A \cdot * (A > \gamma)$$

where  $\cdot *$  is again the pointwise multiplication. This algorithm sets to zero all values of  $A$  under  $\gamma$ , increasing the sparseness of the matrix and consequently the velocity of resolution of the system (4.1); it is evident that this reconstruction is faster but less accurate.

To choose the parameter  $\gamma$  we can proceed in different ways: we can put  $\gamma = \epsilon$ , the machine precision, if we want to delete only the items nearest to zero and not decrease the accuracy of the solution. More typically we choose a fraction of the maximum value of  $A$  such as  $\gamma = 0.05 \max\{a_{i,j}\}$ .



# Chapter 5

## Kernel methods



n alternative choice of the basis in the interpolation can lead to different results in terms of time of computation and accuracy of the solution. Let us recall the interpolation method. Given a basis of functions  $\{b_i(\bar{x})\}_{i=1\dots n}$  that interpolate the function  $f$ , i.e. such that

$$f(\bar{x}) = \sum_{i=0}^n c_i b_i(\bar{x})$$

then for the linearity of the Radon transform

$$\mathcal{R}f(\bar{y}) = \sum_{i=0}^n c_i \mathcal{R}b_i(\bar{y}) \quad \forall \bar{y} \in Y = \{(t_j, \theta_j)\}.$$

We will see in this chapter the interpolation problem solved through the use of kernel functions as basis.

### 5.1 Direct kernel interpolation

To introduce the new basis we need some definitions.

**Definition 8.** A function  $f : \Omega \subseteq \mathbb{R}^s \rightarrow \mathbb{R}$  is called *essentially compact supported* if the closure of the set  $S = \{\bar{x} \text{ s.t. } |B(\bar{x})| > \varepsilon\}$  is compact<sup>1</sup> in  $\mathbb{R}^s$ .

**Remark:** if we choose  $\varepsilon = 0$  the definition above is the classical definition of compact supported function.

---

<sup>1</sup>with  $\varepsilon$  chosen properly, e.g. the machine precision.

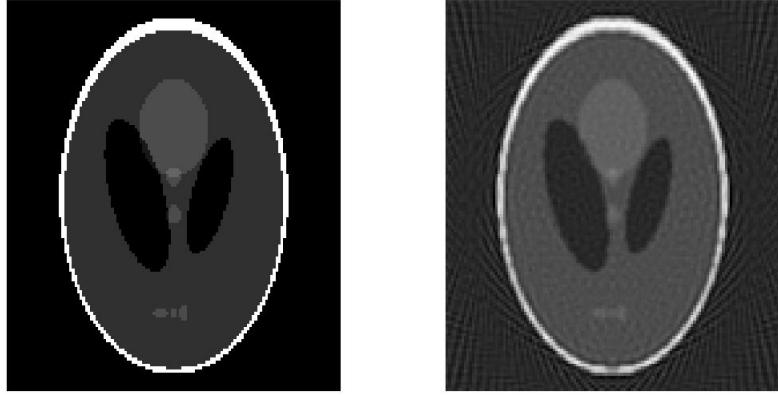


Figure 5.1: Original phantom and kernel reconstruction with Gaussian kernel and shape parameter  $\varepsilon = 1$  after 50 iterations of LSCG.

**Definition 9.** A function  $f : \Omega \subseteq \mathbb{R}^s \rightarrow \mathbb{R}$  is called **radial** (in this work we will also call them **Radial Basis Functions**, briefly or **RBF**) if

$$\phi(\bar{x}) = \varphi(\|\bar{x}\|)$$

where  $\varphi : [0, +\infty] \rightarrow \mathbb{R}$ .

A function  $K : \Omega \times \Omega \subseteq \mathbb{R}^s \times \mathbb{R}^s \rightarrow \mathbb{R}$  is called **Kernel** if

$$K(x, y) = \phi(x - y)$$

with  $\phi$  a radial function on  $\Omega \subseteq \mathbb{R}^s$ .

Instead of the natural pixel basis we use another basis of functions  $b_i(\bar{x}) = B(\bar{x} - \bar{x}_i)$ , where  $B$  is essentially compact supported and radial. For Lemma (3.1.4) the Radon transform of a radial function does not depend from the angular variable  $\theta$ , so to build the matrix we only have to compute  $\mathcal{R}B(\tilde{t}_j)$  where  $\tilde{t}_j = d(\ell_{t_j, \theta_j}, \bar{x}_i)$ , with  $d$  the euclidean distance. Since the line  $\ell_{t, \theta} = (\bar{x}(s)) = (t\bar{\theta} + s\bar{\theta}^\perp)$  has the equivalent, nonparametric form  $t = x \cos \theta + y \sin \theta$ , so the distance from the point  $\bar{x} = (x, y)$  to the line  $\ell_{t, \theta}$  will be

$$\tilde{t}_j = d(\ell_{t_j, \theta_j}, \bar{x}_i) = |x \cos \theta + y \sin \theta - t|$$

In Table 5.1 we show the transform of some RBFs, calculated in order to use them in interpolation. Since the matrix of the interpolation system will be very large, we have chosen (essentially) compact support functions, so that the system matrix will be sparse and the solution of the system will

be computed using one of the iterative algorithms described in the previous chapter.

Function name	$f$	$\mathcal{R}f$
Ball	$\chi_{B(\frac{1}{\varepsilon}, 0)}(r)$	$\sqrt{(\frac{1}{\varepsilon^2} - t^2)}_+$
Gaussian	$e^{-\varepsilon^2 r^2}$	$\frac{\sqrt{\pi}}{\varepsilon} e^{-\varepsilon^2 t^2}$
Wendland $\varphi_{2,0}$	$(1 - \varepsilon r)_+^2$	$\{2\varepsilon^2 \sqrt{\cdot} t^2 + \frac{2}{3}\varepsilon^2 \sqrt{\cdot}^3 - \varepsilon \log(2\frac{1+\varepsilon\sqrt{\cdot}}{1-\varepsilon\sqrt{\cdot}}) t^2\} \chi_{[-\varepsilon, \varepsilon]}$
Wu $\psi_{1,1}$	$(1 - \varepsilon r)_+^2 (\varepsilon r + 2)$	$\{4\sqrt{\cdot} - 3\varepsilon^2 \sqrt{\cdot} - \frac{3}{2}\varepsilon \log(2\frac{\varepsilon+\sqrt{\cdot}}{\varepsilon-\sqrt{\cdot}}) t^2 + \frac{5}{4}\varepsilon^4 \sqrt{\cdot} t^2 + \frac{1}{2}\varepsilon^4 \sqrt{\cdot}^3 + \frac{3}{8}\varepsilon^3 \log(2\frac{\varepsilon+\sqrt{\cdot}}{\varepsilon-\sqrt{\cdot}}) t^4\} \chi_{[-\varepsilon, \varepsilon]}$

Table 5.1: Radon transform of the four chosen functions.

The shorthand  $\sqrt{\cdot} = \sqrt{(\frac{1}{\varepsilon^2} - t^2)}_+$  so that  $t \in (0, \frac{1}{\varepsilon})$  and  $r = \|x\|$ , as usual.

## 5.2 Definite positivity

In the theory of interpolation with kernel functions, which can be found in [4], a typical request is made on the chosen kernel, made to ensure that the interpolation matrix will be nonsingular and positive definite. This condition is just called positive definiteness.

**Definition 10.** A symmetric kernel function  $K$  is called **positive definite** if  $\forall X = x_1, \dots, x_N \subset \Omega$  data points set and  $\bar{c} \in \mathbb{R}^N$

$$\sum_{i,j=1}^N c_i c_j K(x_i, x_j) \geq 0$$

and it's called **strictly positive definite** if the inequality above is verified and the equality stands only for  $\bar{c} = \bar{0}$ .

It is immediate to observe that the matrix of interpolation of a strictly definite positive kernel is definite positive (semi-definite if the kernel is only definite positive).

In our setting the system matrix is not  $K(x_i, x_j)$ , but  $\mathcal{R}(K(x, x_j))_{(y_i)}$ . So we have to verify that the Radon transform of the chosen function is positive definite. Our next step is a theorem that proves that the Radon transform preserves the definite positivity of a radial function. The next theorem will be useful for this purpose.

**Theorem 5.2.1.** Suppose  $\phi \in L^1(\mathbb{R}^d)$  and a continuous function. Then  $\phi$  is positive definite on  $\mathbb{R}^d$  if and only if is bounded and its  $d$ -dimensional Fourier transform  $\mathcal{F}_d(\phi)$  is nonnegative and not identically zero.

*Proof.* see [15] Ch 3 for further details on this theorem.  $\square$

Using this result we prove the following Theorem, which is fundamental in our applications.

**Theorem 5.2.2.** If  $\phi(x - y) = K(x, y)$  is a radial function,  $\phi \in L^1(\mathbb{R}^d)$ , continuous, bounded and positive definite on  $\mathbb{R}^2$ , then its Radon transform  $\mathcal{R}f(t)$  is bounded and positive definite on  $\mathbb{R}^1$ , provided  $\mathcal{R}f \in L^1(\mathbb{R})$ .

*Proof.* A direct consequence of Theorem (5.2.1) is that  $\mathcal{F}_2(\phi)(\bar{\nu})$  is nonnegative and not vanishing, and this property remains if we consider the function in polar coordinates  $\mathcal{F}_2(\phi)(T \cos \theta, T \sin \theta)$ . By means of the Central Slice Theorem (3.3.2), also the function

$$\mathcal{F}_1(\mathcal{R}f)(T, \theta) = \mathcal{F}_2(\phi)(T \cos \theta, T \sin \theta)$$

is nonnegative and not vanishing. Now we remember that the Radon transform of a radial function depends only on the linear variable  $t$ , then

$$\mathcal{F}_1(\mathcal{R}f)(T) \quad \text{is nonnegative and not vanishing}$$

and, again using Theorem (5.2.1) we have the thesis.  $\square$

All functions we are using are positive definite on  $\mathbb{R}^2$ , and so will be their respective transforms. The only exception is the ball function, which is only conditionally positive definite<sup>2</sup> of order 1; we have included that function only to compare the results with other kernels, but we have no warranty about the nonsingularity of the system matrix.

### 5.3 Error estimate

We want to estimate the reconstruction error in case of direct kernel interpolation; if we call such error

$$e(\bar{x}) = f(\bar{x}) - \sum_{i=0}^n c_i K_i(\bar{x})$$

---

<sup>2</sup>a weaker condition, see [4] for the definition.

then for linearity of the Radon transform

$$e(\bar{x}) = \mathcal{R}^{-1}(\mathcal{R}e)(\bar{x}) = \mathcal{R}^{-1} \left( \mathcal{R}f(\bar{y}) - \sum_{i=0}^n c_i \mathcal{R}K_i(\bar{y}) \right)(\bar{x})$$

and so

$$\|e\|_{L^\infty(\mathbb{R}^2)} \leq \|\mathcal{R}^{-1}\|_{(\infty, \infty)} \|\mathcal{R}f - \sum_{i=0}^n c_i \mathcal{R}K_i\|_{L^\infty(\mathbb{R} \times S^1)}$$

where, as in the matrix case,

$$\|\mathcal{R}^{-1}\|_{(\infty, \infty)} = \sup_{g \in L^\infty} \frac{\|\mathcal{R}^{-1}g\|_\infty}{\|g\|_\infty}$$

Now, if we want to proceed with the error estimate we need to show some elements of the approximation theory in interpolation with kernel functions. For a more detailed panoramic about RBF interpolation, see [4] in particular Ch 14.

**Definition 11.** *Given a set  $X$  of  $N$  pairwise distinct points of  $\Omega \subset \mathbb{R}^s$  we call **fill-distance** the radius of the largest ball in  $\Omega$  that does not contain any point of  $X$ , i.e.*

$$h = h_{X, \Omega} = \sup_{x \in \Omega} \min_{x_j \in X} \|x - x_j\|_2$$

**Definition 12.** *Let  $H$  be an Hilbert space  $H = \{f : \Omega \subseteq \mathbb{R}^s \rightarrow \mathbb{R}\}$  with scalar product  $\langle \cdot, \cdot \rangle$ . A Kernel  $K : \Omega \times \Omega \rightarrow \mathbb{R}$  is called **reproducing kernel** for  $H$  if:*

1.  $K(\cdot, x) \in H \quad \forall x \in \Omega$
2.  $f(x) = \langle f, K(\cdot, x) \rangle \quad \forall f \in H, x \in \Omega$

**Definition 13.** *Given  $K$  radial basis function on  $\mathbb{R}^s$  strictly positive definite its **native space** is the Hilbert space  $\mathcal{N}_K(\Omega) = \overline{H_K(\Omega)}$ , where  $H_K(\Omega) = \text{span}\{K(\cdot, y) \mid y \in \Omega\}$  is a pre-Hilbert space with  $K$  its reproducing kernel. Its scalar product is the bilinear form*

$$\left\langle f(\cdot) = \sum_{i=1}^N c_i K(\cdot, x_i), g(\cdot) = \sum_{j=1}^N j_i K(\cdot, x_j) \right\rangle_K = \sum_{i,j=1}^N c_i c_j K(x_i, x_j)$$

given a set  $X$  of  $N$  pairwise distinct points of  $\Omega$ .

**Theorem 5.3.1.** Let  $\Omega \subset \mathbb{R}^s$  be an open, bounded set which satisfies the interior cone condition<sup>3</sup>. Suppose  $K \in C^{2k}(\Omega \times \Omega)$  is a symmetric and strictly positive definite kernel. Denote the interpolant to  $f \in \mathcal{N}_K(\Omega)$  on the set  $X$  of  $N$  pairwise distinct points of  $\Omega$  as  $\mathcal{P}_f$ . Then there exist positive constants  $h_0$  and  $C$  such that, if  $h_{X,\Omega} \leq h_0$ , then

$$|f(\bar{x}) - \mathcal{P}_f(\bar{x})| \leq Ch_{X,\Omega} \sqrt{C_k(\bar{x})} \|f\|_{\mathcal{N}_K(\Omega)}$$

where  $C_k(\bar{x}) = \max_{|\beta|=2k} \max_{w,z \in \Omega \cap B(\bar{x}, c_2 h_{X,\Omega})} |D^{(\beta)} K(w, z)|$ .

**Corollary 5.3.2.** With the same hypothesis as above, there exist positive constants  $h_0$  and  $\tilde{C}$  such that, if  $h_{X,\Omega} \leq h_0$ , then

$$\|f - \mathcal{P}_f\|_\infty \leq \tilde{C} h_{X,\Omega} \|f\|_{\mathcal{N}_K(\Omega)}$$

Now we can formulate the theorem that will give an estimate of the error on the image reconstruction problem.

**Theorem 5.3.3.** Let  $f \in C_0^\infty(B(0, 1))$  a  $b$ -band-limited function, and let  $g = \mathcal{R}f$  be reliably sampled. Let  $K$  be the interpolating Kernel function such that  $\mathcal{R}K$  and its domain satisfy all the conditions of the Theorem (5.3.1). If we call the interpolation error

$$e(\bar{x}) = f(\bar{x}) - \sum_{i=0}^n c_i K_i(\bar{x})$$

then there exist positive constants  $h_0$  and  $\tilde{C}$  such that, if  $h_{X,\Omega} \leq h_0$ , then

$$\|e\|_{L^\infty} \leq 2|S^1| \|w_{ideal}\|_{L^1} \tilde{C} h_{X,\Omega} \|\mathcal{R}f\|_{\mathcal{N}_{\mathcal{R}K}(\Omega)}$$

where  $\|w_{ideal}\|_{L^1} \leq b \sqrt{\frac{N}{18}}$ .

*Proof.* As we have seen before

$$\|e\|_{L^\infty(\mathbb{R}^2)} \leq \|\mathcal{R}^{-1}\|_{(\infty, \infty)} \|\mathcal{R}f - \sum_{i=0}^n c_i \mathcal{R}K_i\|_{L^\infty(\mathbb{R} \times S^1)}$$

where the first part of the second term is

$$\|\mathcal{R}^{-1}\|_{(\infty, \infty)} = \sup_{g \in L^\infty} \frac{\|\mathcal{R}^{-1}g\|_\infty}{\|g\|_\infty}$$

---

<sup>3</sup>or more simply  $\partial\Omega$  has regularity at least  $C^1$ .

and the second one is simply an error of the interpolation using a positive definite kernel.

So all we only have to estimate the term  $\|\mathcal{R}^{-1}\|$  using known results for analytical reconstruction.

From Theorem (3.4.3) we get

$$\|f\|_{L^\infty(\mathbb{R}^2)} \leq 2|S^1| \|w_b\|_{L^1(\mathbb{R})} \|g\|_{L^\infty(\mathbb{R} \times S^1)} + |e_3|.$$

In our formula the transform of the kernel function is made analytically, so the term  $e_3$  is null, and instead of  $\|w_b\|_{L^1}$  we have to use the ideal filter  $w = w_{ideal}$  such that  $\mathcal{R}^*w = \delta$ , the Dirac function. Let's think about the known inversion formulas for the Radon transform.

$$f = \mathcal{R}^* \left[ \mathcal{F}^{-1} \left( \frac{|\nu|}{2} \mathcal{F}(\mathcal{R}f) \right) \right] = \mathcal{R}^* [w * \mathcal{R}f]$$

Let us call  $\Delta = w * \mathcal{R}f$ ; then the Fourier transform gives

$$\mathcal{F}(\Delta) = \mathcal{F}(w)\mathcal{F}(\mathcal{R}f)$$

from one side. On the other side

$$\mathcal{F}(\Delta) = \frac{|\nu|}{2} \mathcal{F}(\mathcal{R}f)$$

So, in conclusion,

$$w = \mathcal{F}^{-1} \left( \frac{|\nu|}{2} \right)$$

The second term would be divergent, but remembering that we are considering a  $b$ -band-limited function, so that the filter must be limited by the radius of acquisition  $r$  and by  $b$  in the frequency domain<sup>4</sup>

$$\|w\|_{L^1(\mathbb{R})} = \|w\|_{L^1([-r, r])} \leq \sqrt{2r} \|w\|_{L^2([-r, r])}$$

using the Parseval's identity on the equivalent form  $\|g\|_2 = \frac{1}{\sqrt{2\pi}} \|\mathcal{F}^{-1}g\|_2$  and knowing that the radius of acquisition<sup>5</sup>  $r = N \frac{\pi}{b}$  with  $N$  as usual the number of pixel per dimension we get

$$\|w\|_{L^1(\mathbb{R})} \leq \sqrt{\frac{2\pi N}{3b}} \cdot \sqrt{\frac{1}{2\pi}} \left\| \frac{|\nu|}{2} \right\|_{L^2([-b, b])} = \sqrt{\frac{N}{12b}} \cdot \sqrt{\frac{2b^3}{3}} = b \sqrt{\frac{N}{18}}$$

---

<sup>4</sup>the inequality is obtained using a known embedding theorem of  $L^p$  spaces on limited domain, if  $0 \leq p < q \leq +\infty$  and the domain of  $f \in L^p$  is a limited set  $S$   $\|f\|_p \leq \text{meas}(S)^{\frac{1}{p} - \frac{1}{q}} \|f\|_q$ .

<sup>5</sup>remember that the size of a pixel is  $\frac{FWHM}{3} = \frac{2\pi}{3b}$ .

This means that

$$\|\mathcal{R}^{-1}\|_{(\infty, \infty)} = 2|S^1| \|w\|_{L^1(\mathbb{R})} \leq 2|S^1|b\sqrt{\frac{N}{18}}.$$

In conclusion

$$\begin{aligned} \|e\|_{L^\infty(\mathbb{R}^2)} &\leq \|\mathcal{R}^{-1}\|_{(\infty, \infty)} \|\mathcal{R}f - \sum_{i=0}^n c_i \mathcal{R}K_i\|_{L^\infty(\mathbb{R} \times S^1)} \leq \\ &\leq 2|S^1| \|w_{ideal}\|_{L^1} \tilde{C} h_{X, \Omega} \|\mathcal{R}f\|_{\mathcal{N}_{\mathcal{R}K}(\Omega)} \end{aligned}$$

which is the promised result.  $\square$

# Chapter 6

## The emission tomography case

Denise and Steffie stayed home that week as men in Mylex suits and respirator masks made systematic sweeps of the building with infrared detecting and measuring equipment. Because Mylex is itself a suspect material, the results tended to be ambiguous and a second round of more rigorous detection had to be scheduled.

D. DeLillo, *White Noise*



As we have said in the previous chapters, for reconstructing an image in the case of emission tomography we will have to deal with attenuation factors. Two different approaches can be used. The first one is to reconstruct the image with the method known in CT, using an attenuation correction method (more used in the PET), with the help of some algorithm for the segmentation and the detection of the contour of the region of interest (ROI).

The second consists of inverting the attenuated Radon transform once we have estimated the attenuation coefficients  $a(x, y)$  from a SPECT/CT hybrid machine (or possibly SPECT/MRI<sup>1</sup>, with greater difficulty). In this chapter we will discuss about the attenuation correction methods and the analytical and iterative algorithms for the inversion of the attenuated Radon transform.

### 6.1 Geometric methods for attenuation correction

The resolution of the reconstruction problem in the case of emission tomography type is complicated because we do not really know the attenuation

---

<sup>1</sup>MRI, Magnetic resonance imaging, another technique used in radiology.

coefficients of the different tissues within the body. To overcome this problem we use two algorithms of attenuation correction which are based on a geometrical estimate of the coefficients.

### 6.1.1 Soreson algorithm

Soreson proposed a pre-correction of the data obtained from the projections. The data is in fact multiplied by the correction factors in order to obtain for each angle the number of counts that would be obtained without attenuation. It is hypothesized the uniform distribution of the tracer and the uniform density of the tissue along each ray; this means that the algorithm will give out good results only when the area to reconstruct is sufficiently uniform, for example the brain. Let  $P_{k,\theta}$  and  $P_{k,\theta+\pi}$  two opposite projections,  $T$  the diameter body and  $t$  the diameter of the active region along the straight line projection.

Then, if we call  $N = \sqrt{P_{k,\theta} \cdot P_{k,\theta+\pi}}$ , the geometric mean of the data the correct value for both angles is

$$N = N_0 e^{\gamma \frac{T}{2}} \left( \frac{f \gamma \frac{T}{2}}{\sinh(f \gamma \frac{T}{2})} \right)$$

where  $\gamma$  is the linear attenuation coefficient (e.g. the brain, if the tracer used and  $^{99m}T$  we suppose  $\gamma = 0,035 \text{ cm}^{-1}$ ) and  $f = \frac{t}{T}$  is a factor which in practice varies from 0.5 to 0.75. We can fix that value, for instance to 0.6 or better to an average estimated value of  $\frac{t}{T}$ .

### 6.1.2 Chang algorithm

This algorithm makes a post-correction on the reconstructed image. The result of the classic reconstruction is multiplied pixel by pixel by the weights obtained from an estimate of the mean attenuation that the beam with angle  $\theta$  had been subject to at the transition from pixel  $(x, y)$ .

Then, if  $P$  is the image obtained e.g. by filtered back projection and  $P_0$  the corrected image, then  $P_0(x, y) = c(x, y)P(x, y)$ , where the correction terms are:

$$C(x, y) = \left( \frac{1}{l} \sum_{i=1}^l e^{-\mu l_{\theta_i}} \right)^{-1}$$

where  $l_{\theta_i}$  is the distance between the pixel  $(x, y)$  and the edge of the body at the angle  $\theta_i$  and  $\mu$  is the linear attenuation coefficient, supposed to be constant.

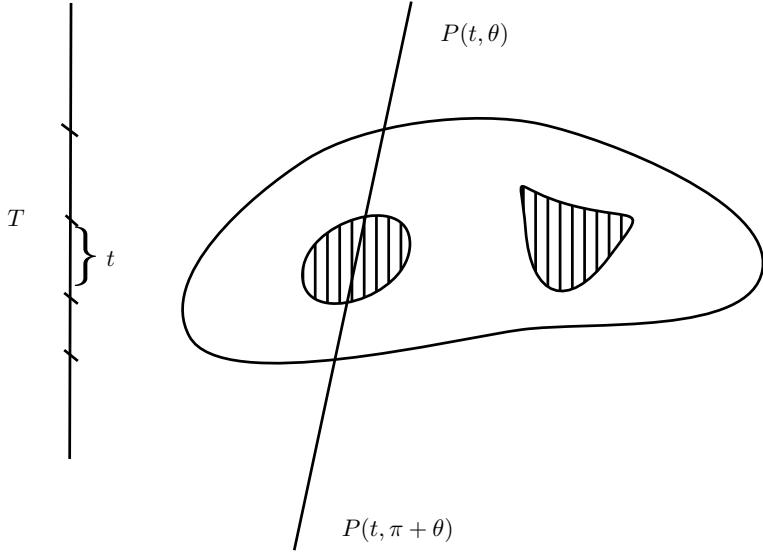


Figure 6.1: Scheme of the data used by the algorithm of Soreson.

Therefore also the Chang algorithm gives good results if the tested area is uniform<sup>2</sup>.

## 6.2 Inversion formula of the attenuated Radon transform

We recall the AtRT formula

$$\mathcal{R}_a f(t, \theta) = \int_{\ell_{(t, \theta)}} e^{-\mathcal{D}a(\bar{x}, \theta + \frac{\pi}{2})} f(\bar{x}) \quad (6.1)$$

now, given  $a(\bar{x})$  we want to recover the value of  $f$  from its transform  $g = \mathcal{R}_a f$ . The first analytical solution to this problem was given in [14] by G. Novikov in 2000. We will see the version of this formula published in [12] by F. Natterer in the same year.

First of all we need the definition the Hilbert transform.

**Definition 14.** Let  $g$  a suitable function, then its **Hilbert transform** is the function

$$\mathcal{H}g(s) = \frac{1}{\pi} \int_{\mathbb{R}} \frac{g(t)}{s - t} dt \quad (6.2)$$

where the integral is meant as a Cauchy principal value.

---

<sup>2</sup>for a more detailed explanation, see the original article [1].

Note that this transform can be thought as the convolution  $g * h$  where  $h(s) = \frac{1}{\pi s}$ , but the convolution integral does not converge.

Now let us define the function

$$h = \frac{1}{2}(I + i\mathcal{H})\mathcal{R}a \quad (6.3)$$

where  $I$  is the identity operator,  $i$  the imaginary unit, and both the operator are meant to be applied on the linear variable of  $\mathcal{R}a$ .

Now we can enunciate following Theorem

**Theorem 6.2.1. (Novikov-Natterer formula)** *Let  $g = \mathcal{R}_a f$  for  $f$  transformable function, and  $h$  the function (6.3). Assume  $a(x, y)$  known, then  $f$  is uniquely determined by the following formula*

$$f(\bar{x}) = \frac{1}{4\pi} \Re \operatorname{div} \int_{S^1} \theta e^{\mathcal{D}a(\bar{x}, \theta + \frac{\pi}{2})} (e^{-h} \mathcal{H} e^h g)_{(\bar{x}, \bar{\theta}, \theta)} d\theta \quad (6.4)$$

where  $S^1 = [0, 2\pi]$ .

*Proof.* see [12] for the complete proof.  $\square$

Now let  $a = 0$  everywhere, then  $g = \mathcal{R}f$ ,  $h = 0$  and  $\mathcal{D}a = 0$  so that

$$f(\bar{x}) = \frac{1}{4\pi} \int_{S^1} (\mathcal{H}g')_{(\bar{x}, \bar{\theta}, \theta)} d\theta = \mathcal{R}^*(\mathcal{H} \frac{\partial}{\partial s} g)$$

which is an equivalent form<sup>3</sup> of (3.3).

### 6.3 Discretization of the AtRT inversion formula

The discretization<sup>4</sup> of the (6.4) is similar to the discretization of the FPB formula (3.4). In fact all we have to do is to take the “attenuated back-projection” of the convolution product between  $g$  and the inverse Fourier transform of a low-pass filter.

First of all we have to compute the function  $g_a$

$$g_a = \Re(e^{-h} \mathcal{H} e^h g)$$

Putting  $h = h_1 + ih_2$ ,  $h_1 = \frac{1}{2}\mathcal{R}a$  and  $h_2 = \frac{1}{2}\mathcal{H}\mathcal{R}a$  then

$$g_a = e^{-h_1} (\cos h_1 \mathcal{H} e^{h_1} \cos h_2 + \sin h_2 \mathcal{H} e^{h_1} \sin h_2) g$$

---

<sup>3</sup>the proof is in [5] Ch 6 p.51.

<sup>4</sup>see [12] Ch 3, p.6 for more details about the discretization.

where, like in the non-attenuated case,  $\mathcal{H}$  is approximated by a convolution. Let  $\sigma_b$  a low pass filter<sup>5</sup> with  $b$  its cut-off frequency and  $\Sigma_b = \mathcal{F}^{-1}\sigma_b$ , then

$$\mathcal{H}_b f = \Sigma_b * f$$

once we have calculated  $g_a$  we only have to compute

$$f(\bar{x}) = \frac{1}{4\pi} \operatorname{div} \int_{S^1} \theta e^{\mathcal{D}a(\bar{x}, \theta + \frac{\pi}{2})} g_a(\bar{x} \cdot \bar{\theta}, \theta) d\theta = \mathcal{R}_a^*(g_a),$$

where the integral is computed in the sense of a discrete mean and the divergence operator is approximated by finite differences as explained in [12].

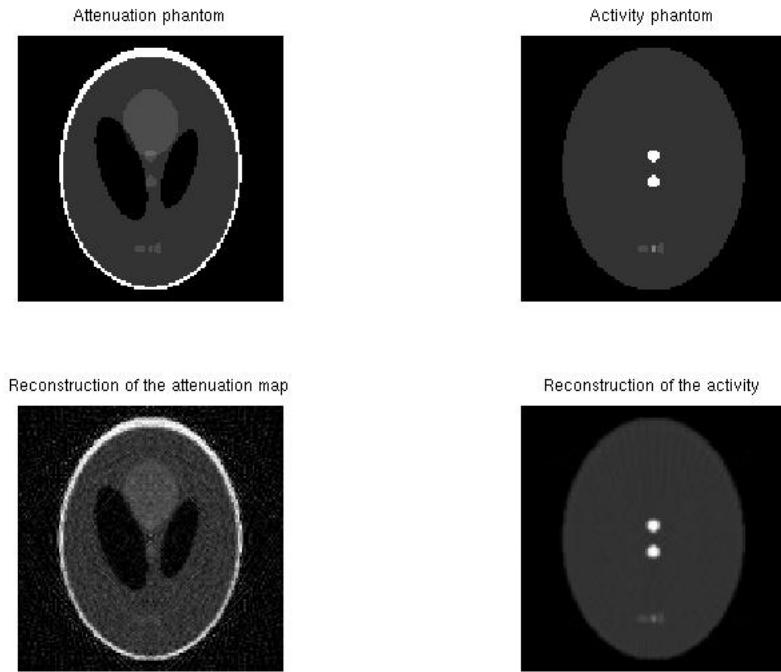


Figure 6.2: Analytical reconstruction of a SPECT/CT phantom data.

## 6.4 Iterative methods

We propose a new method of reconstruction for the attenuated case. The main idea is the use of iterative methods as in the transmission tomography.

---

<sup>5</sup>again, for a list of the low-pass filter used see the chapter about regularization.

We consider the outgoing Gamma-rays at angle  $\theta$  coming from a pixel with unit concentration of tracer.

As we said in previous chapters, the AtRT is not equivalent to the RT of a weighted function. Nevertheless from the matrix  $A$  used in the transmission case, we can obtain a matrix  $B$  such that  $Bc = d$  where  $B_{i,j} = \mathcal{R}_a b_i(t_j, \theta_j)$ ,  $c_i = f(\bar{x}_i)$  the unknown term and  $d_j = \mathcal{R}_a f(t_j, \theta_j)$  the data.

Let us consider, as before, the basis  $b_i(\bar{x}) = \chi_{P_i}$  and assume

$$f(\bar{x}) = \sum_{i=0}^n c_i b_i(\bar{x}).$$

Then for linearity

$$\mathcal{R}_a f(\bar{y}) = \sum_{i=0}^n c_i \mathcal{R}_a b_i(\bar{y}) \quad \forall \bar{y} \in Y = \{(t_j, \theta_j)\}$$

Now to compute the elements of  $B$ , we have to remind that the data of attenuation come from a CT reconstruction, i.e., as suggested in [7], are on the form

$$a(\bar{x}) = \sum_{k=1}^{N^2} g_k \chi_{P_k}(\bar{x}).$$

For linearity its fan beam transform will be

$$\mathcal{D}a(\bar{x}, \theta) = \sum_{k=1}^{N^2} g_k \int_0^{+\infty} \chi_{P_k}(\bar{x} + t\bar{\theta}) dt = \sum_{k=1}^{N^2} g_k \text{meas}(P_k \cap \ell_{\bar{x}, \theta}^+)$$

where  $\ell_{\bar{x}, \theta}^+ = \{\bar{x} + t\bar{\theta} \mid s \in \mathbb{R}^+\}$  and  $\text{meas}$  is the Lebesgue measure. Now let us consider the AtRT of the basis  $b_i$

$$\mathcal{R}_a b_i = \int_{\ell_{t, \theta}} \exp \left( - \sum_{k=1}^{N^2} g_k \text{meas}(P_k \cap \ell_{\bar{x}, \theta + \frac{\pi}{2}}^+) \right) \chi_{P_i}.$$

This last formula is hard to compute, but it offers a useful interpretation. In fact if the tracer is concentrated only on the  $i$ -th pixel  $P_i$ , the outgoing rays will be subjected to an attenuation equal to the attenuated sum of the path length; if  $I_{in} = 1$ , according to Beer-Lambert's law

$$I_{out} = I_{in} \exp \left( - \sum_{k=1}^{N^2} g_k \text{meas}(P_k \cap \ell_{\bar{x}, \theta}^+) \right).$$

Now, if we consider the matrix  $A$  used in CT tomography, we can compute the outgoing rays from the pixel  $P_i$  in  $(t_j, \theta_j)$  as

$$B_{i,j} = A_{i,j} \exp \left( - \sum_{(k_1, k_2) \in K_{(i,j)}} g_k \text{ meas}(P_k \cap \ell_{t_j, \theta_j}) \right) = \\ A_{i,j} \exp \left( - \sum_{(k_1, k_2) \in K_{(i,j)}} g_k A_{k_1, k_2} \right)$$

where  $K_{(i,j)} = \{(k_1, k_2)\} \subset \{1, \dots, N^2\}^2$  is the set of the indexes of the pixels covered by the line  $\ell_{x_i, \theta_j}^+$  and the index  $k = lp - ((k_1 - 1)p + k_2) + 1$ .

This method requires a computation time a little longer than that of the

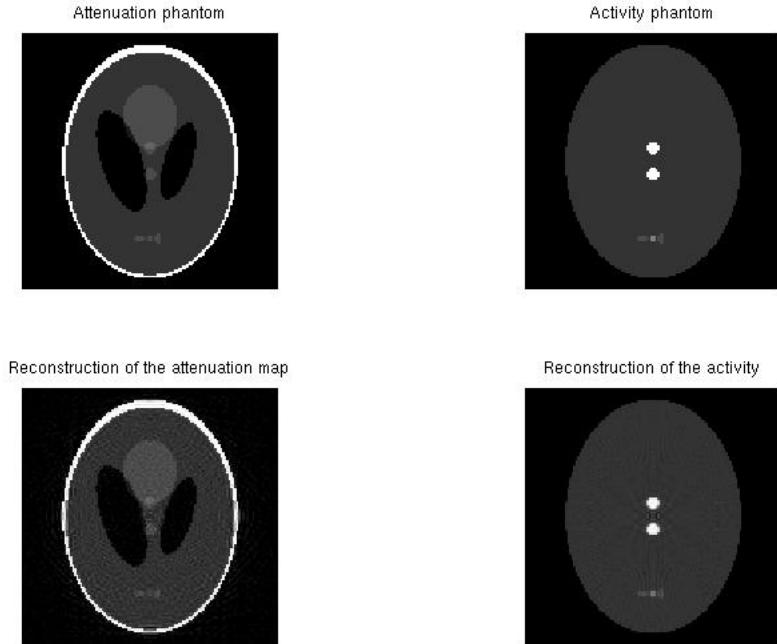


Figure 6.3: Iterative reconstruction of a SPECT/CT phantom data with  $\lambda = 0.1$ .

analytical approach because the system matrix  $B$  is calculated *a posteriori*. Moreover this matrix has often a large condition number, so the analytical solution seems to be preferable this time.

Nevertheless, a regularization of the linear system can bring to a more accurate solution than the analytically reconstructed image. First of all we can

introduce a relaxation parameter  $\lambda \in [0, 1]$  in order to consider relatively the effect of the attenuation

$$B_{i,j}^{(\lambda)} = A_{i,j} \exp \left( -\lambda \sum_{(k_1, k_2) \in K_{(i,j)}} g_k A_{k_1, k_2} \right).$$

Note that  $B^{(0)} = A$  and  $B^{(1)} = B$ . With this little foresight the linear system can lead to a more accurate solution.

Regularization methods will be introduced in the next chapter.

# Chapter 7

## Regularization methods and post-processing

he matrix of the linear systems to solve both in transmission and emission cases is large, sparse, and often ill-conditioned. It can even happen that the matrix had pair of identical columns and some null columns. Moreover, the scattering can bring to a noisy data. So it is reasonable to add a regularization in the reconstruction which will bring an image with less artifacts<sup>1</sup>. We also introduce a contour-detection algorithm.

### 7.1 Low-pass filters

Let  $L$  be the cut-off frequency e.g. the resolution frequency  $b$  found in equation (3.2). Now we will choose the filter  $w_L$ , and convolve its Fourier transform with the reconstructed image  $P$ . The result will be a filtered version of  $P$ .

$$\tilde{P} = \mathcal{F}^{-1}[w_L \mathcal{F}(P)] = P * \mathcal{F}^{-1}w_L$$

The following functions are commonly used in filtering medical images:

- “Ram-Lack” filter

$$w_L(|\|\nu\||) = \chi_{[-L,L]}(|\|\nu\||)$$

---

<sup>1</sup>a regularization method in general increases the smoothness of the solution, i.e. in our case increases the partial volume effect and the number of artifacts can increase. So we will always use regularization methods in a further step, instead of embedding it into the reconstruction process. With this expedient we will have two images to compare: an unfiltered, raw image with better resolution and a filtered, smooth image, but with a lower resolution.

which is simply the truncation of frequencies larger or smaller than  $L$ .

- “Shepp-Logan” filter

$$w_L(\|\nu\|) = \left( \frac{\sin(\frac{\pi\|\nu\|}{2L})}{\frac{\pi\|\nu\|}{2L}} \right) \chi_{[-L,L]}(\|\nu\|)$$

essentially a truncated sinc function.

- “low-pass cosine” filter

$$w_L(\|\nu\|) = \cos\left(\frac{\pi\|\nu\|}{2L}\right) \chi_{[-L,L]}(\|\nu\|)$$

a normalized, truncated cosine function.

In our notation we wrote  $\|\nu\|$  not as a norm but as  $\|\nu\| = \max\{\nu_1, \nu_2\}$ . The functions used as filters in FBP formula and its analogue in the emission case are exactly the version of those filters in  $\mathbb{R}^1$  (then we use  $\nu$  instead of  $\|\nu\|$ ) multiplied by the absolute value of  $\nu$ .

$$\omega_L(\nu) = |\nu| w_L(\nu)$$

If we want to find the discrete value of the cutoff frequency  $L$  we have to know the following proportion  $L : f_s = cut : N$  where  $cut$  is the discrete cutoff,  $N$  the numbers of samples and  $f_s$  the sampling frequency, which for a standard dimension problem is  $f_s = 1/3.5\text{mm}^{-1} = 0.286\text{mm}^{-1}$  because 3.5 mm is the typical value for the physical length of a pixel<sup>2</sup>.

The results of filtering in our experiments are not encouraging. In fact filtering does not bring to a better image, either in terms of  $\infty$ -norm error or in terms of visual quality.

## 7.2 Philips-Tichonov regularization

Philips-Tichonov is a classic regularization method used for ill-conditioned system, its main purpose is to add regularity to the solution; the typical iterative algorithm for solving linear systems minimizes the quantity

$$\min_{x \in \mathbb{R}^n} \|Ax - b\|_2$$

---

<sup>2</sup>more precisely the length of a pixel is FWHM/3, and the typical value for FWHM is about 10mm around 150mm of distance.

now we add a penalty term

$$\min_{x \in \mathbb{R}^n} \|Ax - b\|_2 + \gamma \|Fx\|_2$$

where  $\gamma$  is a positive, tunable constant, and  $F = \Delta^m$  the regularization matrix that makes the solution similar to a function of regularity  $C^m$ . Typically  $\Delta_{i,i} = 1$  and  $\Delta_{i,j} = -1$  if the  $i$ -th pixel is next to the  $j$ -th, zero otherwise<sup>3</sup>. The solution of the new system will be

$$(A^t A + \gamma F^t F)x = A^t b$$

and we will find this solution iteratively.

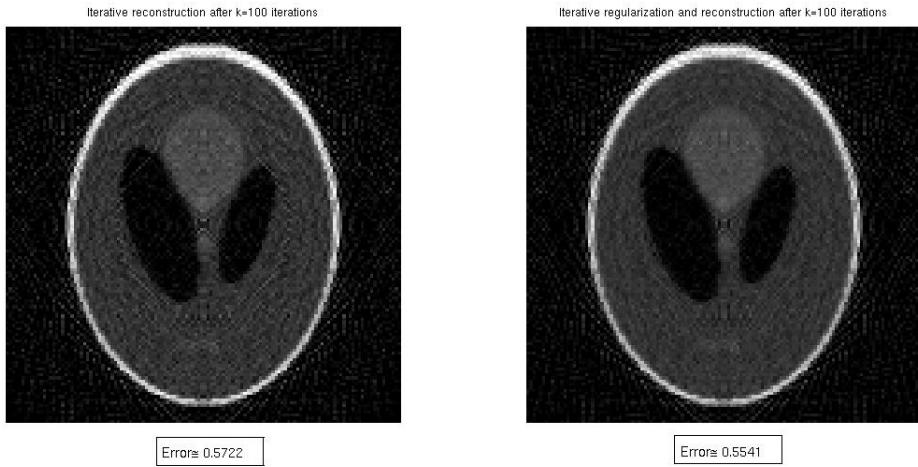


Figure 7.1: Reconstructed image (left) and its regularization (right). As we can see from the  $\infty$ -norm errors a regularization does not always bring a great improvement to the quality of the solution.

### 7.3 A segmentation algorithm

A **phantom** is a function to test the efficiency of the algorithm *in silico*<sup>4</sup>, i.e. without data given by a physical machine. It is also called phantom a physical device that simulates a certain organ and used to make experiments on the machine. For clarity we will call the second one **dummy**.

<sup>3</sup>we assume  $\Delta^0 = I$ , the identity matrix.

<sup>4</sup>this term is uses commonly in biomedical engineering; it means that experiments are not made on a living been (*in vivo*) or on a few component or chemical reproductions (*in vitro*), but only on a mathematical model.

Usually a phantom is a step function, or rather  $f = \sum_i \chi_{B_i}$  with  $B_i \subset \mathbb{R}^2$  some compact sets. This type of functions is congenial to the reconstruction because it splits the region of interest (or ROI) in different areas where the attenuation coefficient is constant. We can assign constant that if we recognize the type of tissue or calculate in the case of SPECT/CT. Therefore we need a segmentation algorithm that levels the image obtained by the CT image reconstruction, making it like a phantom, which we call phantomization algorithm. The proposed algorithm is the following:

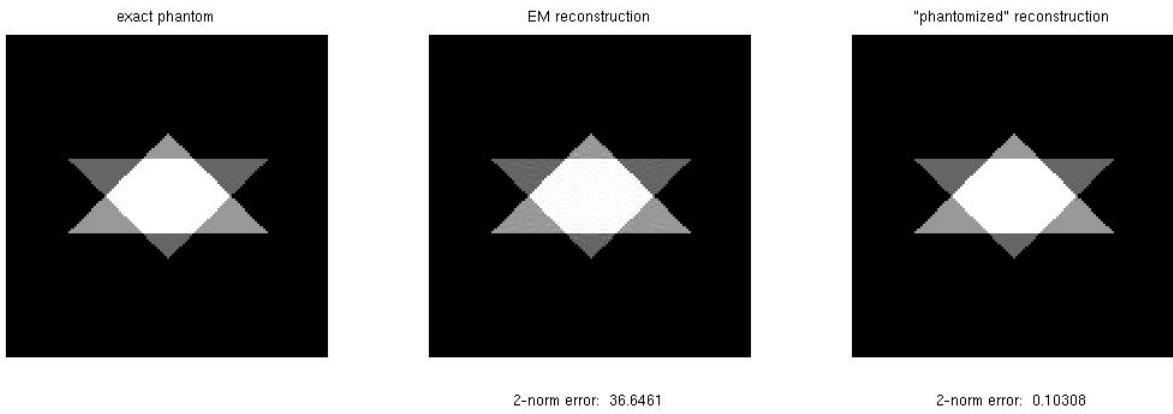


Figure 7.2: Reconstruction of a David star phantom and its phantomization; as we can see from the errors the phantomized solution is close to a phantom.

1.  $P$  is the  $N \times N$  matrix of reconstructed image,  $k$  the number of different levels that you want to look in  $[0.1]$
2. put its elements in a column vector,  $P = P(:)$  long  $N^2$
3. for  $i = 1, \dots, N^2 - 2$ 
  - (a) we calculate  $S$  the sum and  $D$  the difference of the elements in  $X = P(i : i + 2)$
  - (b) if  $S$  or  $D$  has norm less than  $\frac{2}{k}$  then  $P(i : i + 2) = \text{mean}(X)$  were  $\text{mean}$  is the the average
  - (c) otherwise  $P$  remains the same
  - (d) end of the *for* loop
4. repeat 2. and 3. for each rotation of  $\frac{\pi}{2}$  of the matrix  $P$

This algorithm, if the starting reconstruction is sufficiently accurate, returns an image very similar to a phantom. We can alternatively assign to  $P$  different values as  $P(i : i + 2) = 1/k * \text{round}(k * \text{mean}(X))$  where  $\text{round}$  is a rounding function. In this way the zones are more uniform, but the value of the attenuation coefficient tends to be less accurate unless we choose  $k$  *ad hoc* for the values of attenuation coefficients in the examined area; e.g. if we know that the coefficients are all integral multiples of  $\frac{1}{10}$ . For instance  $k = 10$  is a good value.

## 7.4 A simple algorithm for the contour detection of the ROI

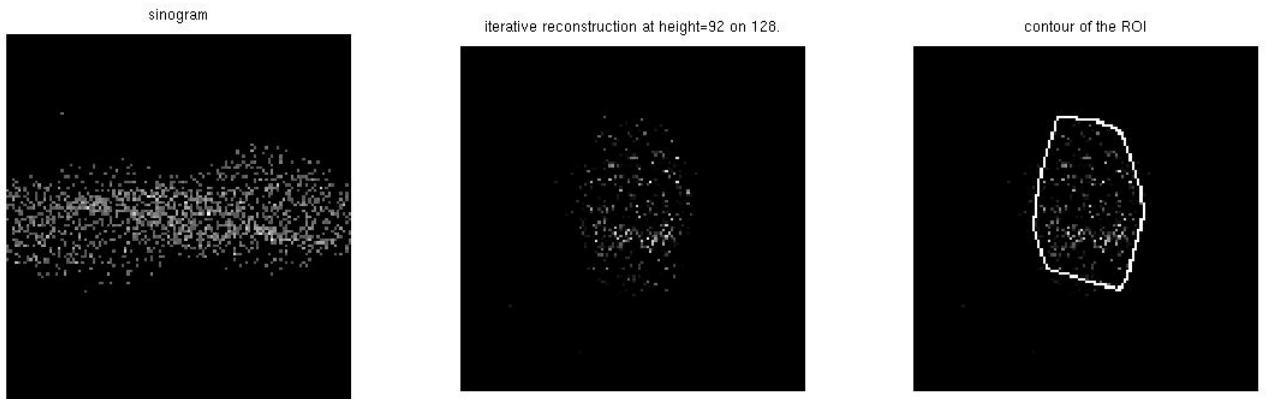


Figure 7.3: Edge detection of a brain dummy's data; the algorithm detects efficiently the ROI's contour.

There are many algorithms written for the detection of the contours of an image, the one we are proposing is a very simple application of the convex hull algorithm of a point set.

Let  $x$  be the solution e.g. from an iterative reconstruction. Now we identify the area in which is the traced is more concentrated with a simple zero-padding the solution  $x = x.*(x > \delta)$  where  $\delta = \gamma \max(x)$  and  $\gamma \in (0, 1)$  is chosen properly (e.g. we have chosen  $\gamma = 0.22$  in our experiments). Now we apply a convex hull algorithm 'convhull' and draw the segments between the external points.

The following MATLAB<sup>®</sup> function builds a matrix  $B$  with all zeros except

for the contour points, where the matrix has value 1.

```
function B=edge_detection(I,rapp)
%home-made edge detection for SPECT reconstructed images

B=zeros(128,128);
m=max(I)*rapp;
II=zeropadding(I,m);
II=reshape(II,128,[]);
[x,y]=find(II>=m);
if length(x)>2
k=convhull(x,y);

lung=30; %lung is chosen properly to represent every pixel of the line
t=linspace(0,1,lung);
for i=2:length(k)
    a=round(t.*x(k(i-1))+(1-t).*x(k(i)));
    b=round(t.*y(k(i-1))+(1-t).*y(k(i)));
    for j=1:lung
        B(a(j),b(j))=1;
    end;
end;
end;
```

# Chapter 8

## Experimental results from *in silico* tests

When I was a child, I used to talk as a child, think as a child, reason as a child; when I became a man, I put aside childish things.

At present we see indistinctly, as in a mirror, but then face to face. At present I know partially; then I shall know fully, as I am fully known.

1Cor, 13: 11-12



e now present some experimental results from several tests made on phantoms (*in silico*). The aim of these experiments is to test the algorithms presented in this work and to compare the quality of the respective reconstructed images. All the tests has been made on the standard dimensions, i.e. we have reconstructed a 128x128 Shepp-Logan phantom by taking 3°-step angular projections of 128 linear data each, except where is differently specified, and the error will be evaluated in  $\infty$ -norm, coherently with the error estimates made in previous chapters.

The computer used for all these tests have a 3.5 Gb of RAM equipped with two dual-processors of 2.70 GHz each on a 32-bit architecture.

### 8.1 Filtered Back Projection results

We now proceed by evaluating the performs of the FBP algorithm in terms of precision and time of computation. In these experiments we are using the standard *radon* and *iradon* functions included in the image processing toolbox package of MATLAB®, which standards are different: it reconstructs

128x128 images by a 185x60 matrix data. Now, to test the error bound provided by Theorem (3.4.3), which is

$$\|f - \tilde{f}\|_{L^\infty(\mathbb{R}^2)} \leq 2|S^1| \ \|w_b\|_{L^1(\mathbb{R})} \|g - \tilde{g}\|_{L^\infty(\mathbb{R} \times S^1)} + |e_3|,$$

we have to compute the  $L^1$ -norm of the filters in use.

- “Ram-Lack” filter

$$\|w_L(\nu)\|_{L^1} = 2 \int_0^L \nu \, d\nu = L^2 = c_{(1)} L^2$$

- “Shepp-Logan” filter

$$\|w_L(\nu)\|_{L^1} = 2 \int_0^L \nu \operatorname{sinc}\left(\frac{\pi\|\nu\|}{2L}\right) \, d\nu = \frac{8}{\pi^2} L^2 = c_{(2)} L^2$$

- “low-pass cosine” filter

$$\|w_L(\nu)\|_{L^1} = 2 \int_0^L \nu \cos\left(\frac{\pi\|\nu\|}{2L}\right) \, d\nu = \frac{4(\pi-2)}{\pi^2} L^2 = c_{(3)} L^2$$

All the constants  $c_{(i)}$  are  $O(1)$ , so the different choice of filters does not change the order of the error bound. We choose the constant  $c_3 = 0.3974$  as the FBP error with no noise on the signal.

In Figure 8.1 is shown the error of the FBP algorithm using the “Ram-Lack” filter for several noise errors.

The FBP algorithm is quite fast, its computational cost is  $O(N^2 \log N)$ .

The mean time of computation for this experiment is  $t \cong 0.088$ s. As we can see from Figure 8.2 the FBP can reconstruct quite large images in less than a second.

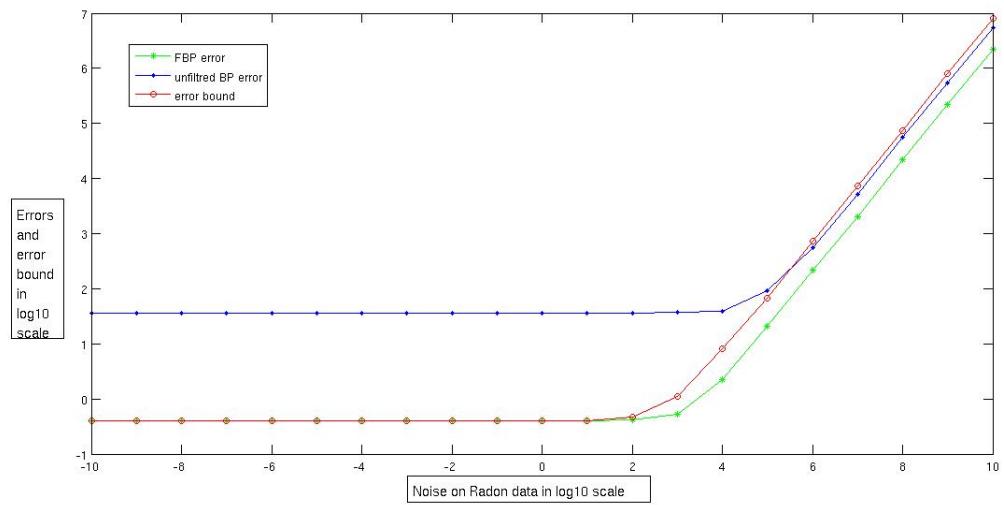


Figure 8.1: A test for the error bound using the standard “Ram-Lack” filter; note that the unfiltered back projection error is bigger than the bound for low noise levels, and reaches the same order of the bound for higher noise levels.

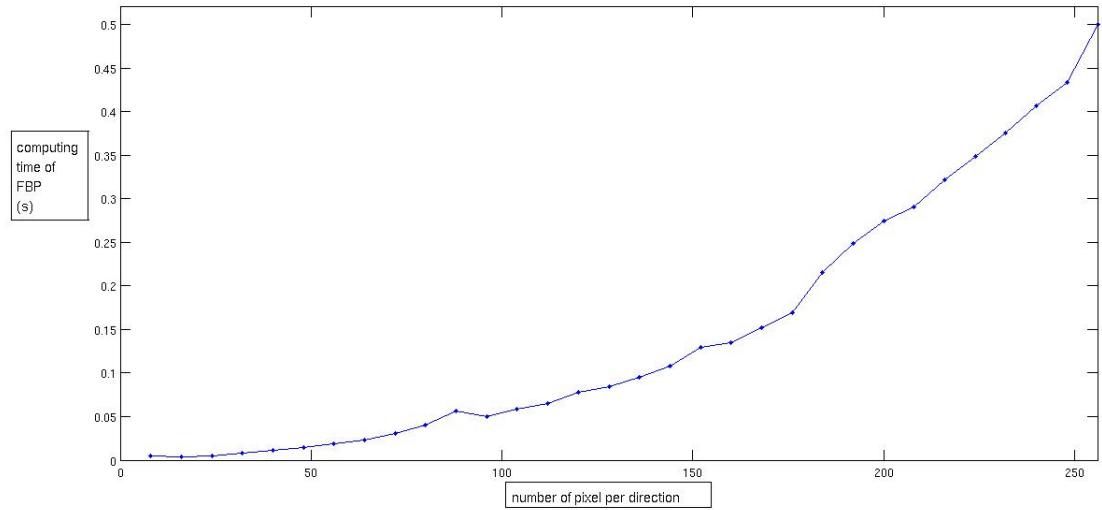


Figure 8.2: Time test for the FBP algorithm.

## 8.2 Iterative Methods results

Next experiments concern the iterative methods described in Chapter 4. Since the ART algorithm is very slow<sup>1</sup> we proceed by showing the results of MLEM and LSCG methods. The algorithms in these experiments are exactly the ones described in the corresponding, sections, while the algorithm used for computing the system matrix  $A$  is contained in the “Air tools package”,[8] available in Hansen’s web page<sup>2</sup>.

In this section, we will show some tests on several iteration numbers and noise levels.

In the case the noise on the data are null, both algorithms show fast and good performances. The results of our test are shown in Figure 8.3. In both cases the time of computation grows linearly with the number of iterations and the error is almost everywhere decreasing. This confirms that the computational cost for  $k$  iterations is  $O(kN^2)$ .

We note that MLEM stops after the 75th iteration. In practice, MLEM algorithm is more precise than the LSCG, but its result could be instable in the case the  $k$ -th iteration gives a value of  $c^f := Ax^{(k)}$  close to zero so that the punctual division  $c^q := c./c^f$  is unaccurate.

We observe from the results that the lowest error for MLEM is  $err \cong 0.4260$  at iteration 700, which is obtained after a time  $t \cong 4.177$  while the lowest error for LSCG is  $err \cong 0.5668$  at iteration 800, which is obtained after a time of  $t \cong 4.162$ .

Now we test our algorithms in the case the Radon data is subject to a Gaussian relative noise of null mean and standard deviation equal to  $\sigma$ . We have chosen the value of  $\sigma = 10\%$  for the first experiment. As we can see in Figure 8.4, that the MLEM algorithm is more accurate but unstable, and the LSCG error reflects the typical behavior of the iterative methods in noisy cases. In fact the error does not decrease by raising the number of iterations, but exists an optimal iteration number, in our case around 200, where the error is minimal. This is reasonable, because the iterative algorithm tries to reach the image whose sinogram is closer to our noisy data, and it will reach a noisy image. The discussion in medical image reconstruction about the choice of the optimal iteration remains open, because it is dependent from the image itself.

As a second experiment, we tried to vary the relative noise level from 0 to

---

<sup>1</sup>it typically shows a good results only after  $N^2$  iterations, i.e. about 16000 for the standard case.

<sup>2</sup><http://www2.imm.dtu.dk/~pch/AIRtools/>.

100%, and use the same number of iteration for both algorithms. In Figure 8.5 we can observe that the MLEM is more accurate, but again, is “allergic” to some noise levels at some iteration and the error of LSCG does not raise so much with noise and can bring good results even if the noise level is high.

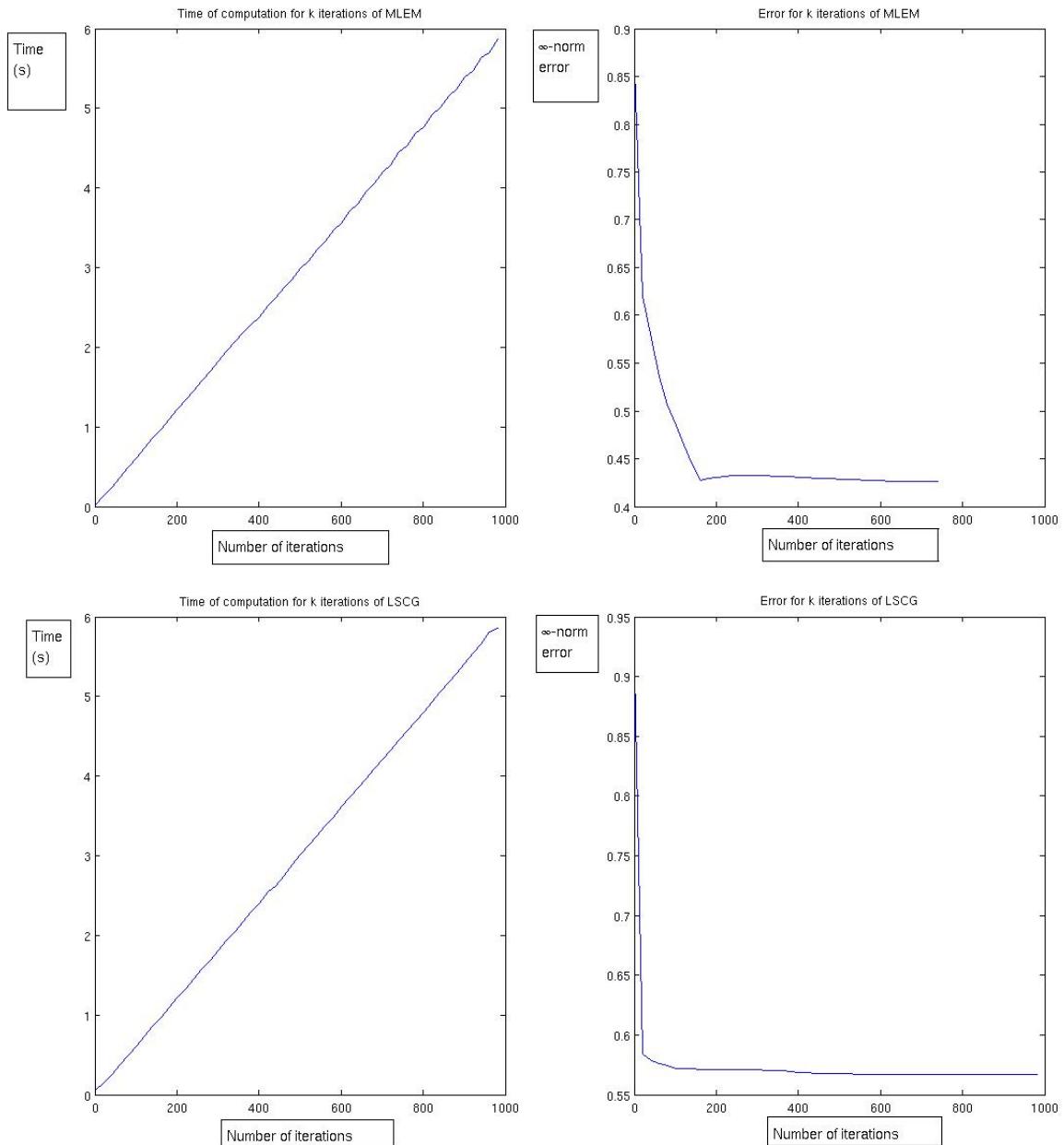


Figure 8.3: Time and error test for MLEM (upper) and LSCG (lower) at several iterations.

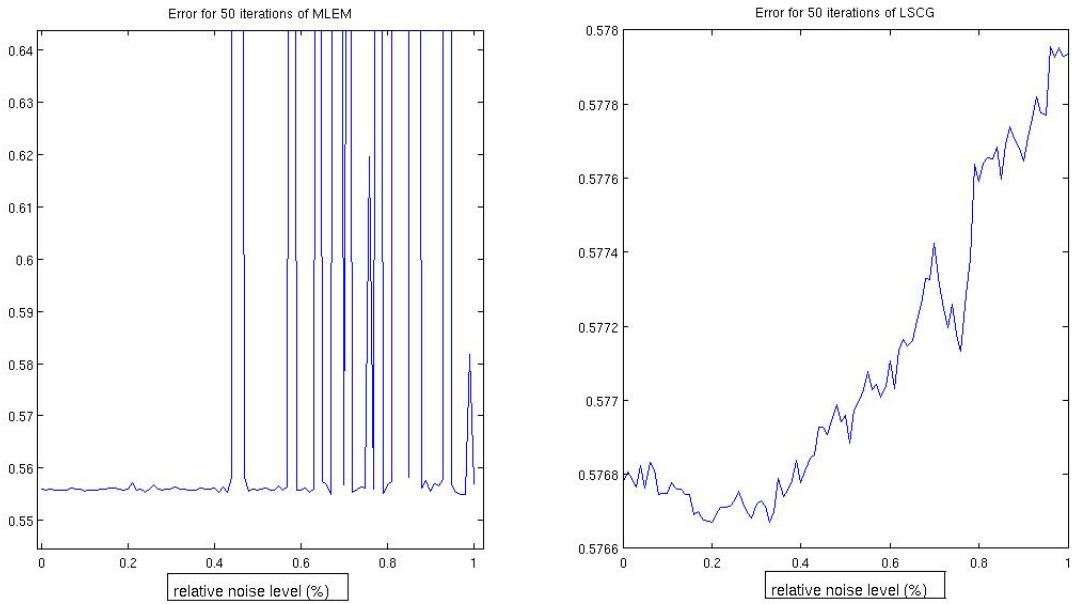


Figure 8.5: Error of the MLEM and LSCG algorithms with several noise levels.

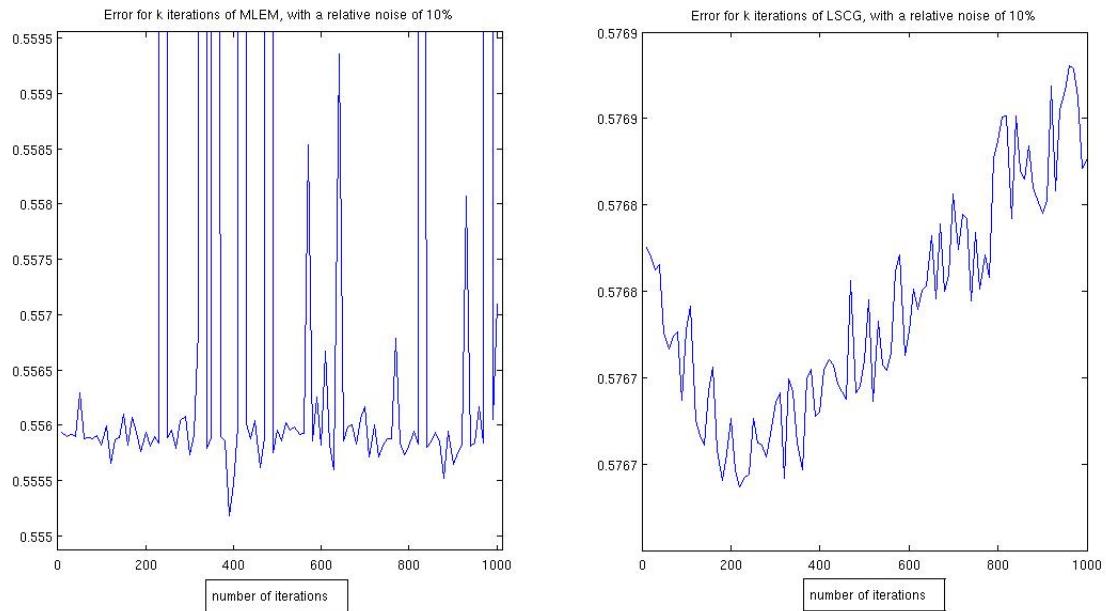


Figure 8.4: Error of the MLEM and LSCG algorithms with a noise of  $\sigma = 10\%$  after several numbers of iterations.

### 8.3 Kernel methods results

We made<sup>3</sup> some tests also on the kernel method described in Chapter 5, using the four functions we have selected before: the Gaussian function the Ball function, the Wendland and the Wu function. As we can see in Figure 8.6, the sparsity of the system matrix depends on the shape parameter  $\varepsilon$ : as  $\varepsilon$  decreases, the support becomes larger and the system matrix becomes less sparse (i.e. it has less elements equal to zero). If the matrix has few nonzero elements the results are inaccurate, while on the other side if the nonzero elements are too much the time needed for assembling the system matrix and the time of solution of the system is very long. Therefore a good trade-off between time computation and accuracy is needed. For this reasons we have chosen to work on values of  $\varepsilon \in [0.5, 10]$ . We made also further studies on a cluster with the aim to investigate the results for smallest values of  $\varepsilon$ . We have chosen for this test to use the LSCG algorithm for 50 iterations because we considered this choice as a good trade-off between speed and accuracy.

We also have found the “optimal” shape parameter with trial-and-error, i.e. we have selected the value of  $\varepsilon$  which gives a lower error. Finally we have verified the error bound proved in Theorem (5.3.3), i.e.

$$\|f - \tilde{f}\|_{L^\infty} \leq 2|S^1| \|w_{ideal}\|_{L^1} \tilde{C} h_{X,\Omega} \|\mathcal{R}f\|_{\mathcal{N}_{RK}(\Omega)}$$

where  $\|w_{ideal}\|_{L^1} \leq b\sqrt{\frac{N}{18}}$ . Now  $|S^1| = \pi$ ,  $b = \frac{2\pi}{118} \cong 0.053$ , then  $\|w_{ideal}\|_{L^1} \leq b\sqrt{\frac{128}{18}} \cong 0.141$  and  $h_{X,\Omega} = \frac{\sqrt{2}}{2}$  because the data centers are equispaced with unitary distance.

We have estimated the constant  $\tilde{C}$  experimentally, as

$$\tilde{C} = \max \frac{\|f - \tilde{f}\|_\infty}{2|S^1|b\sqrt{\frac{N}{18}}h_{X,\Omega}\|\mathcal{R}f\|_{\mathcal{N}_{RK}}}$$

and verified the error bound. We can see the results of these tests in Figures 8.7 and 8.8; the constant  $\tilde{C}$  the optimal shape parameter  $\varepsilon_{opt}$  are in Table 8.1.

---

<sup>3</sup>our scripts are based on the Air Tools package, as all the iterative algorithms in this work.

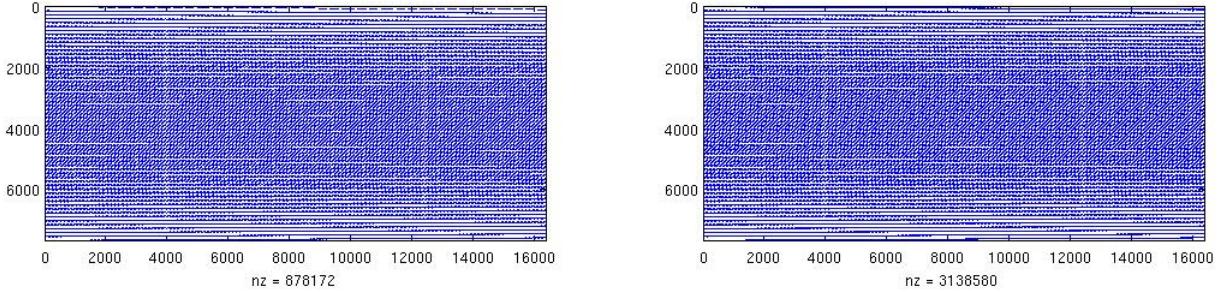


Figure 8.6: Sparsity of the classical ART matrix and of the matrix obtained using Gaussian kernel and  $\varepsilon = 1$ ; the shape parameter  $\varepsilon$  can change the (essential) support of the kernel and the sparsity too. A too sparse matrix can be too ill-conditioned, while a “richer” matrix slows down the reconstruction. A good trade-off is required in the choice of the shape parameter.

function	$\tilde{C}$	$\varepsilon_{opt}$	error for $\varepsilon_{opt}$
Gaussian	0.0406	1.5	0.4903
Ball	0.0437	1	0.4957
Wendland	0.0446	1.5	0.6011
Wu	0.0434	1	0.5104

Table 8.1: This table shows the results of the experiments made on kernel functions for  $\varepsilon \in [0.5, 10]$ ; for each function is reported the corresponding estimate for the constant  $\tilde{C}$ , the optimal shape parameter, and the  $\infty$ -norm error of the corresponding reconstruction.

We made two more tests for  $\varepsilon \in [0.01, 1]$ . As we mentioned before we have used cluster computing for such values of the shape parameter. To be more precise we used the *hpblade14* knot of the NumLab at the Department of Mathematics<sup>4</sup>. This knot uses two processors Quad-Core AMD Opteron<sup>TM</sup>, 2.3GHz, and 64Gb of RAM memory, of which we have used 32Gb.

In Figure 8.9 we see the results of two different sampling of  $\varepsilon$ : the first one is a quasi-equispaced sampling  $\varepsilon \in \{0.01, 0.05, 0.1, 0.15, \dots, 1\}$ , the second one considers different rays of the (essential) support  $\frac{1}{\varepsilon} \in \{1, 2, \dots, 20\}$ . As we can see from both the results, the optimal shape parameter for Ball and Wu functions is still  $\varepsilon_{opt} = 1$ . On the contrary the optimal parameter  $\varepsilon_{opt}$  is smaller for the Gaussian ( $\varepsilon_{opt} = 0.55$ ) and Wendland ( $\varepsilon_{opt} = 0.091$ ).

Concerning the speed, the iterative methods have the same order of convergence than the classic iterative methods,  $O(kN)$ , but are often slower be-

<sup>4</sup>see the page <http://numlab.math.unipd.it/?q=node/23> for further details.

cause they require also the multiplying the coefficients for the data matrix<sup>5</sup>  $H_{i,j} = K(x_i, x_j)$ .

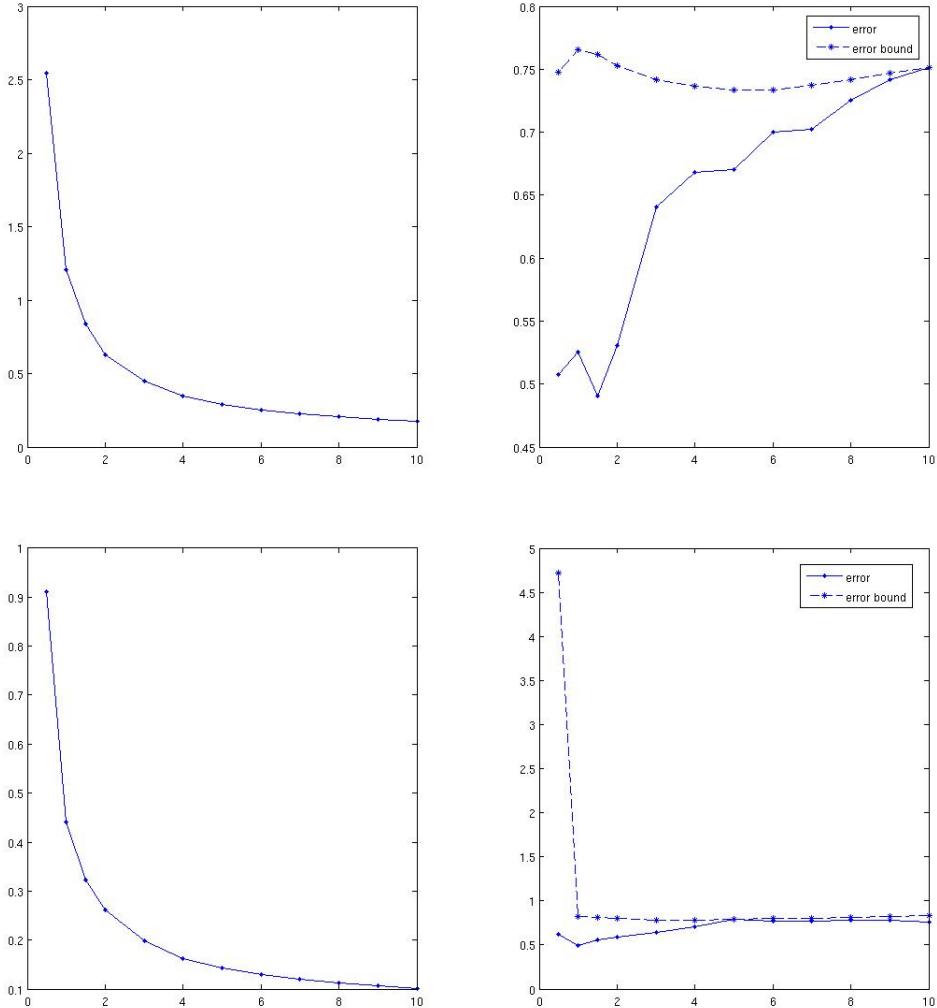


Figure 8.7: Time of computation in seconds (left), error and error bound (right) in  $\infty$ -norm of the kernel method for the functions Gaussian (upper) and "ball" (lower) with several shape parameters  $\varepsilon \in [0.5, 10]$

<sup>5</sup>if the shape parameter is big it happens that  $H = I$  the identity matrix, and the computational time required for solving our system is less than the time of computation for a classical MLEM or LSCG algorithm.

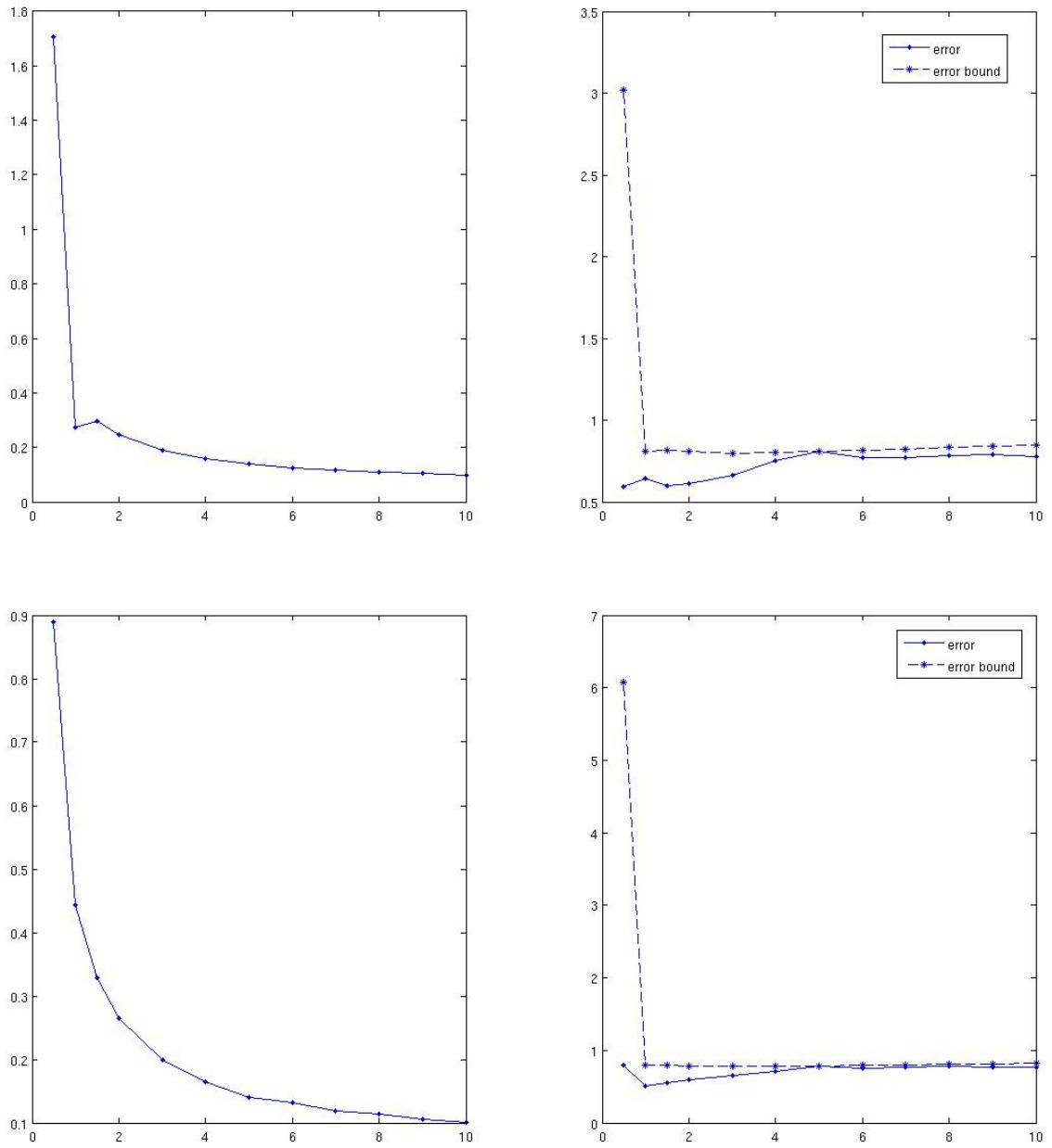


Figure 8.8: Time of computation in seconds (left), error and error bound (right) in  $\infty$ -norm of the kernel method for the functions Wendland (upper) and Wu (lower) with several shape parameters  $\varepsilon \in [0.5, 10]$ .

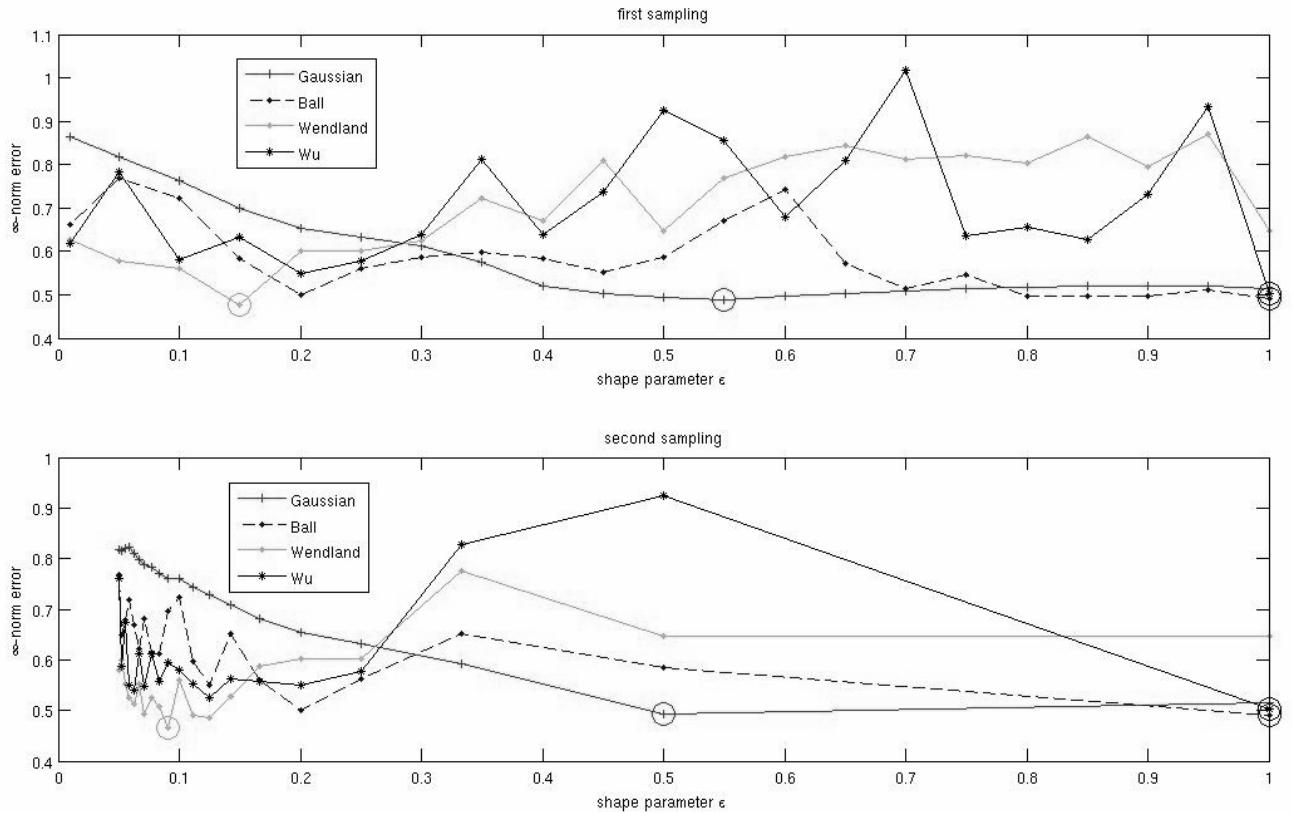


Figure 8.9: Results of two tests made with cluster computer for two different samplings of  $\varepsilon \in [0.01, 1]$ . The optimal shape parameter  $\varepsilon_{opt}$  is signed with a circle.

## 8.4 Comparisons of the results

We will now compare the results of the different algorithms in terms, as usual, of  $\infty$ -norm error.

Let us compare the efficiency of the FBP algorithm<sup>6</sup> with the iterative LSCG algorithm at 50 iterations and the kernel method using Gaussian function, optimal shape parameter and 50 iterations of LSCG method. As we can see in Figure 8.10, both the iterative tests lead to more accurate results than those of the FBP, and the Gaussian can improve the results of the LSCG. If we look at the mean time of computation we see that the FBP runs on  $t_1 \cong 0.1s$ , while the LSCG and Gaussian methods require to be ran at time

<sup>6</sup>when this time we have forced the dimensions so that it reconstruct a 128x128 image from a 128x60 data matrix.

$t_2 \cong 0.3\text{s}$  and  $t_3 \cong 0.8\text{s}$  respectively. This means that the kernel method with Gaussian reduces the error of 15% requiring 160% more time with respect to the LSCG method.

As a second comparative test we will vary the number of iterations and compare the classic iterative methods with the respective kernel versions, i.e. the kernel method using such methods. The comparison will be made on time of computation and error. We have chosen for this test the Gaussian function which appears to be the most precise and slow of the kernels and for the Wendland function, which seems to be the fastest and least accurate. As we see in Figure 8.11, the results obtained are what we expected in LSCG case. In fact the Gaussian is the slowest but most accurate, while the Wendland is the fastest but least accurate, while the classic method offers mean results both in terms of speed and of accuracy. To the contrary the MLEM test offers a surprise: the Gaussian function's results appear to be faster after a few iterations, but the classical method passes the kernel method after 70 iterations. The Wendland function's results remain always the faster but the least accurate between the three methods.

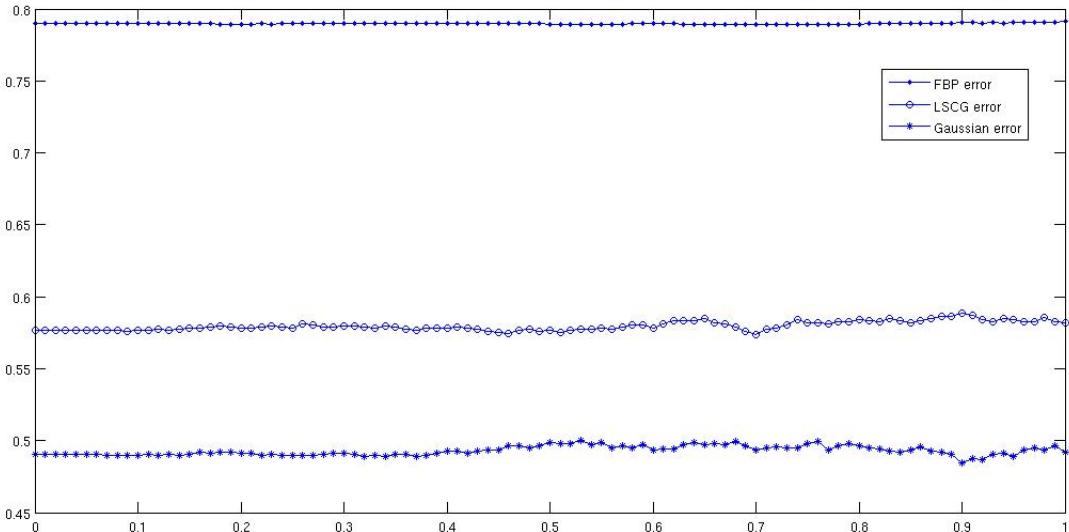


Figure 8.10: Error of FBP, LSCG, and Gaussian algorithms as the relative error varies from 0 to 100%. The results of the FBP formula are worse than those in the previous section because we're forcing the FBP algorithm to work on the same dimensions of the other algorithms and not on its standard dimensions.

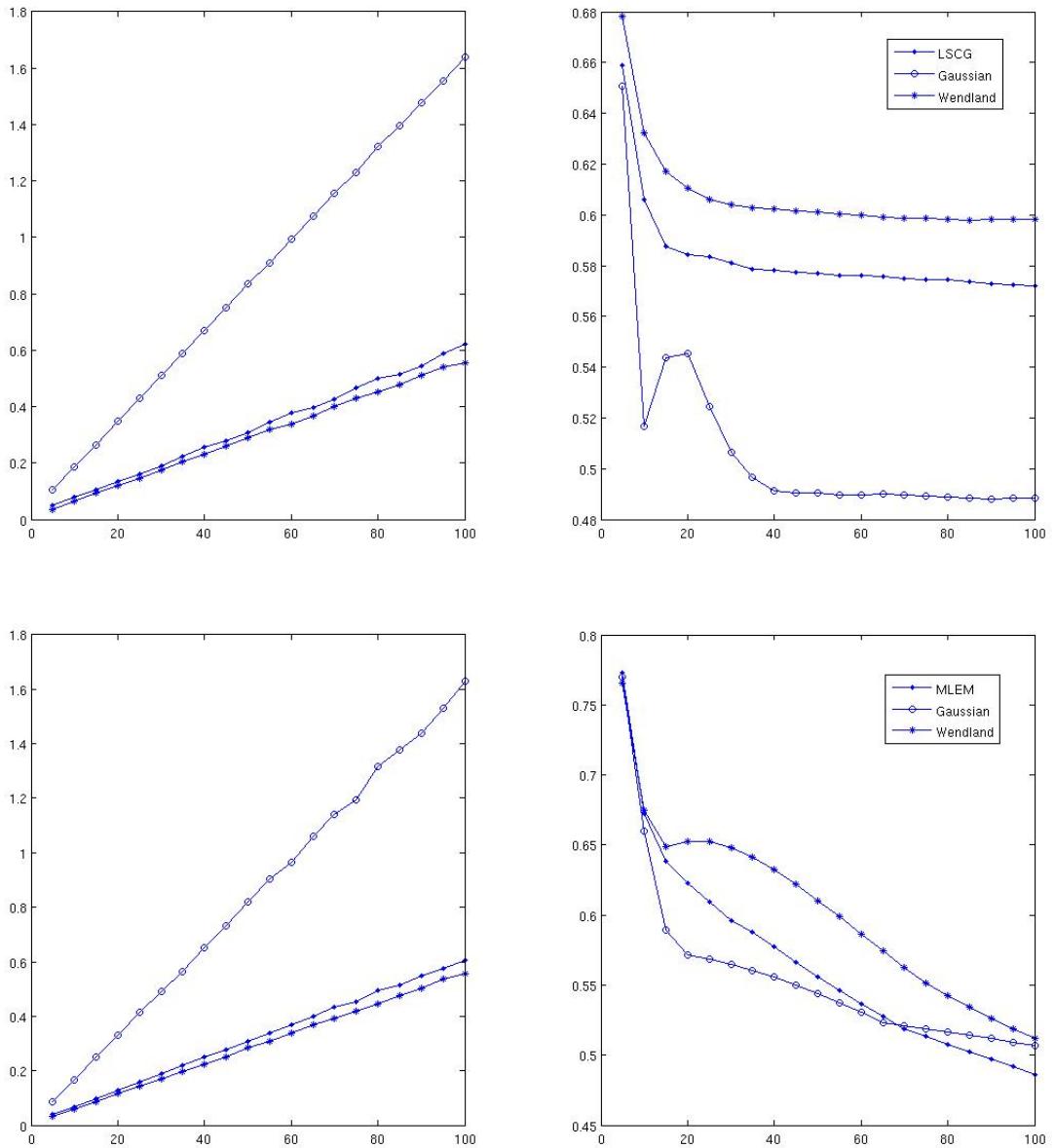


Figure 8.11: Computational times in seconds (left) and errors (right) for the LSCG and its kernel versions (upper), for MLEM and its kernel versions (lower) with Gaussian and Wendland functions. Along x we have the number of iterations.

## 8.5 Hybrid SPECT/CT simulations

In this last section we will compare the results of the two methods we have proposed for the hybrid SPECT/CT reconstruction. The first one is the analytic algorithm based on Novikov-Natter formula<sup>7</sup>, very similar to the FBP, while the second one is based on the iterative methods<sup>8</sup>, introduced in Chapter 6. Since the analytical algorithm does not allow any change of dimension we will work on its standards: we will reconstruct a 128x128 image using a data matrix 182x120 and a (reconstructed) attenuation map 128x128. For this purpose we are using two phantoms as in Figures 6.2 and 6.3 in Chapter 6. The first one is the attenuation map, that will be projected and back-projected using the iterative LSCG at 100 iterations. Then we will use the reconstructed attenuation map to compute the attenuated Radon transform of the second phantom and invert the transform using the two methods.

Before comparing the two methods we will choose the parameter that weights the effect of the attenuation. For this purpose we will simply make some reconstruction (always using 100 iterations of LSCG) at on some values of  $\lambda \in [0, 1]$  and select the value that gives a smallest  $\infty$ -norm error. The results of this test are shown in Figure 8.12. As we can see the error can change very much, while the total time needed to compute the system matrix and to solve the system is almost constant. This test indicates us the value of  $\lambda_{opt} = 0.04$  as the “optimal” value, which gives an error of  $e = 0.235$ .

Now we can compare the two algorithms in terms of  $\infty$ -norm error of the reconstruction of noisy signals. As we have done in previous section we will compare the errors of the two methods when the data are affected from a relative noise which goes from 0 to 100%. For the iterative method we have used again 100 iterations of LSCG algorithm and the optimal value  $\lambda_{opt} = 0.04$  found in the previous experiment.

From Figure 8.13 we can see that the time required for the solution of the system with iterative method is a little longer than the time used by the analytical one. On the other hand the iterative formula reveals a lowest error and a better resistance to the noise.

---

<sup>7</sup>for our tests we have used the functions made by F. Monard and available at his webpage <http://www.math.washington.edu/~fmonard/>.

<sup>8</sup>and then is based on the Air Tools package as all the iterative algorithms in this work.

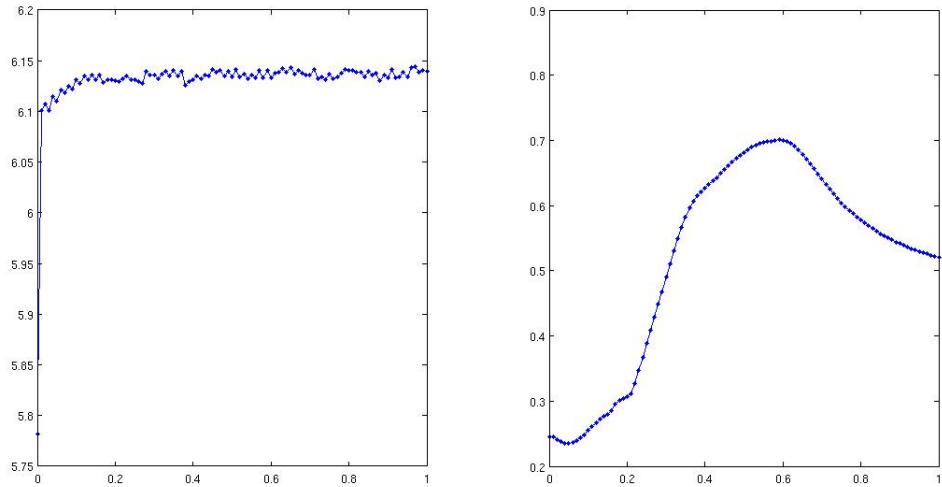


Figure 8.12: Trial-and-error test for the  $\lambda$  parameter. To the left we see the total computational time of the matrix  $B$  and the solution of the system. Right: the errors for various values of  $\lambda \in [0, 1]$ .

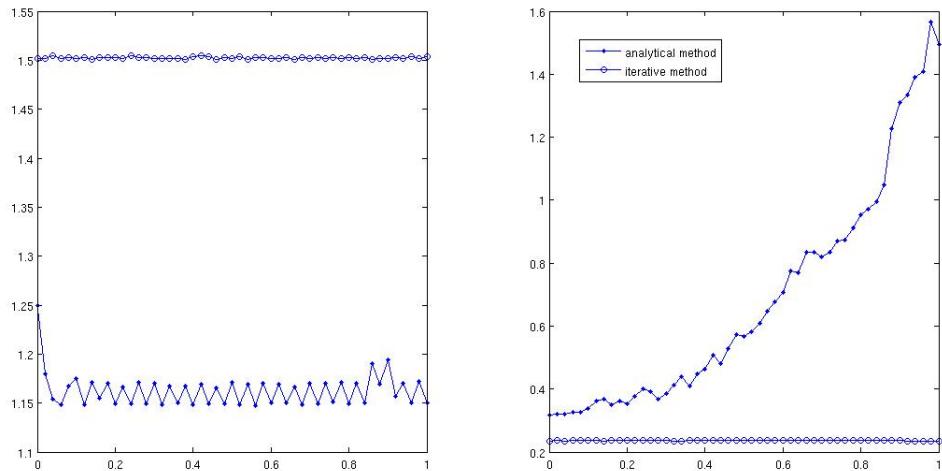


Figure 8.13: Time of resolution in seconds (left) and error (right) for the analytical and iterative methods for the resolution of the hybrid SPECT/CT simulated problem at several relative noise levels from 0 to 100%.



# Chapter 9

## Results from *In vitro* experiments

As far as the laws of mathematics refer to reality, they are not certain; and as far as they are certain, they do not refer to reality.

Albert Einstein, quote reported in the 1960 article *The Unreasonable Effectiveness of Mathematics in the Natural Sciences* by E. Wigner.



This chapter is dedicated to some experiments made on dummies (*in vitro*). The quality of the results is not evaluated scientifically, we will only make some visual considerations.

We will also discuss about the resolution of the Gamma camera we are using, and compare it with the experimental resolution.

### 9.1 Resolution and sampling

Before we discuss about the laboratory results, we will verify that the standard sampling (120x128) is correct with respect to the system resolution and does not give data with lower resolution than the resolution of the machine. First of all we have to compute the machine resolution, i.e. the resolution of the collimator in use in our machine.

The gamma camera used in the experiments is the Ultra High Resolution (UHR), and the collimators have a hexagonal structure, like the hives of bees, and the crystal type is 3/4".

In Table 9.1 are reported the main characteristics of the most commons machines used in SPECT.

Description	Max. useful energy (Kev)	Septal thickness (mm)	Hole size (mm)	Hole length (mm)	Resolution at 0mm (FWHM, mm)	Resolution at 100mm (FWHM, mm)
Crystal type					3/8"; 3/4"	3/8"; 3/4"
Dynamic	150	0.305	2.54	25.4	5.1; 5.7	14.8; 15.0
High Resolution	160	0.203	1.22	27.0	3.9; 4.7	7.0; 7.6
UHR	160	0.152	1.78	58.4	4.1; 4.8	6.0; 6.7
Medium energy	300	0.864	3.40	58.4	4.1; 4.8	6.0; 6.7
High energy	440	2.01	3.4	58.4	5.9; 6.6	11.0; 11.4

Table 9.1: Main characteristics of the most commons machines used in SPECT.

The declared intrinsic resolution is  $R_i = 4.1\text{mm}$ .

Let us take a look on the septal thickness. From our formula we find

$$t \gtrsim \frac{6D}{\mu l - 3} = \frac{6 \cdot 1.78\text{mm}}{2.14\text{mm}^{-1} \cdot 58.4\text{mm} - 3} \cong 0.088\text{mm}$$

The septal thickness of our machine,  $t = 0.152\text{mm}$ , fulfills the condition above.

The collimator resolution is  $R \cong D(1 + \frac{x}{L_{eff}})$  if we consider that the collimators are made of lead,  $\mu = 2.14\text{mm}^{-1}$  at 150 keV, then the effective collimator length is

$$L_{eff} = L - \frac{2}{\mu} \cong 58.4\text{mm} - 0.9\text{mm} \cong 57.5\text{mm}$$

So we can compute the collimator resolution

$$R_c = \frac{1.78}{57.5}x + 1.78 = 0.031x + 1.78$$

This means that the resolution at 0 mm is 1.78mm and at 100 mm we have a resolution of 4.9 mm.

We can also calculate the sensitivity

$$S \cong \left( \frac{KD}{L_{eff}} \right)^2 \cong \left( \frac{0.26 \cdot 1.78\text{mm}}{57.5\text{mm}} \right)^2 \cong 6.47 \cdot 10^{-5}.$$

What is the minimal sampling which guarantees a resolution of  $R_s = 6.7\text{mm}$  (the biggest resolution at 100mm)? Now, this resolution means a maximum frequency at

$$b = \frac{2\pi}{R_s} 100 \cong 93.78$$

so the minimum number of angular samples is  $l = \lfloor b \rfloor + 2 = 95$  and the minimum number of equispaced linear samples is  $p = \lfloor \frac{2b}{\pi} \rfloor = 59$ . This means that the standard sampling is correct and does not decrease the resolution. We can also obtain from this data an estimate of the system resolution, which is  $R_s = \sqrt{R_i^2 + R_c^2}$ , where we know the value of  $R_i = 4.1\text{mm}$ . This means that the system resolution is  $R_s \cong 4.47\text{mm}$  at 0mm and  $R_s \cong 6.8\text{mm}$  at 100m, which is consistent with the values of FWHM declared on Table 9.1.

## 9.2 Three capillaries experiment and experimental resolution

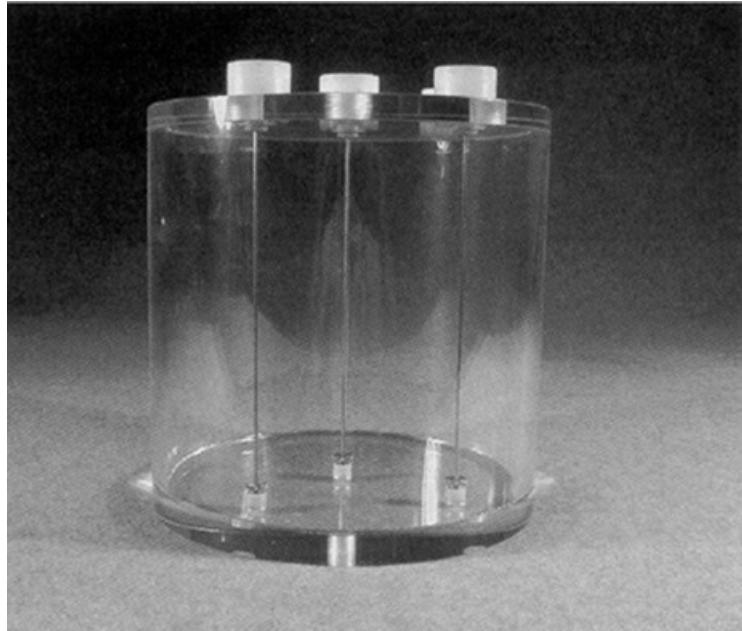


Figure 9.1: The three capillary dummy used in experiments.

The first dummy used in experiments, shown in Figure 9.1, is a cylinder filled with water with three thin tubes on the inside, filled with tracer. This experiment is a valid alternative to the small sphere filled with tracer for the estimation of the experimental resolution as FWHM of the reconstructed

image. The data are taken with different radius of acquiring, precisely  $r_1 = 145\text{mm}$ ,  $r_2 = 165\text{mm}$  and  $r_3 = 185\text{mm}$ . With this data we will be able to extrapolate an estimation of the resolution at a desired distance.

The raw data is a  $128 \times 128 \times 120$  matrix which contains the projection data taken in 120 angular projection by a grid of  $128 \times 128$  collimators. This is equivalent for our purposes to 128 sinograms  $128 \times 120$ , each one of them represents the analysis at different heights.

In Figure 9.2 we can see a reconstruction and segmentation of a slice of this dummy, obtained using the LSCG at 100 iterations and the phantomization algorithm.

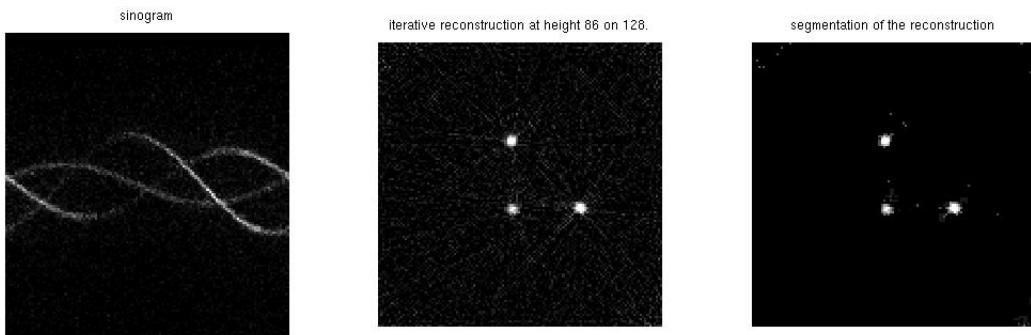


Figure 9.2: Raw data sinogram, reconstruction and phantomization for the three capillaries dummy data.

From the reconstruction of the central capillary at different slices we obtain some estimates of the FWHM<sup>1</sup>. The mean values of those FWHM are reported on Table 9.2.

distance	145mm	165mm	185mm
FWHM	8.78	9.53	10.62

Table 9.2: Values of the experimental FWHM at several distances between abject and collimators.

From these three values we can extrapolate which could be a reasonable experimental resolution at a given distance. In fact the intrinsic resolution  $R_i = \gamma$  is a constant, while the collimator resolution depends linearly from the distance  $R_c = \alpha x + \beta$ , with  $\alpha, \gamma, \beta > 0$ . So the system resolution will be

---

<sup>1</sup>this work has done by prof. Cecchin, we want to thank him again for his fruitful cooperation.

of the form

$$FWHM(x) = R_s = \sqrt{R_c^2 + R_i^2} = \sqrt{\alpha^2 x^2 + 2\alpha\beta x + \beta^2 + \gamma^2} = \sqrt{ax^2 + bx + c}$$

with  $a, b, c > 0$ .

If we want to find the constants  $a, b, c$  we have to solve this problem

$$\min_{a,b,c} \|\sqrt{ax^2 + bx + c} - f\|^2$$

with the constraints  $a, b, c > 0$ . In our notation now  $x = (45, 165, 185)$  the distance vector and  $f = (8.78, 9.53, 10.62)$  are the data, and the chosen norm is the discrete 2-norm.

To solve this problem we have used the function *fmincon* contained in the MATLAB® package *Optimization Toolbox*. The results are shown in Figure 9.3. From these results we find a resolution of 4.71mm at 0mm and of 6.79mm at 100mm, which is close to the declared values of FWHM in Table 9.1.

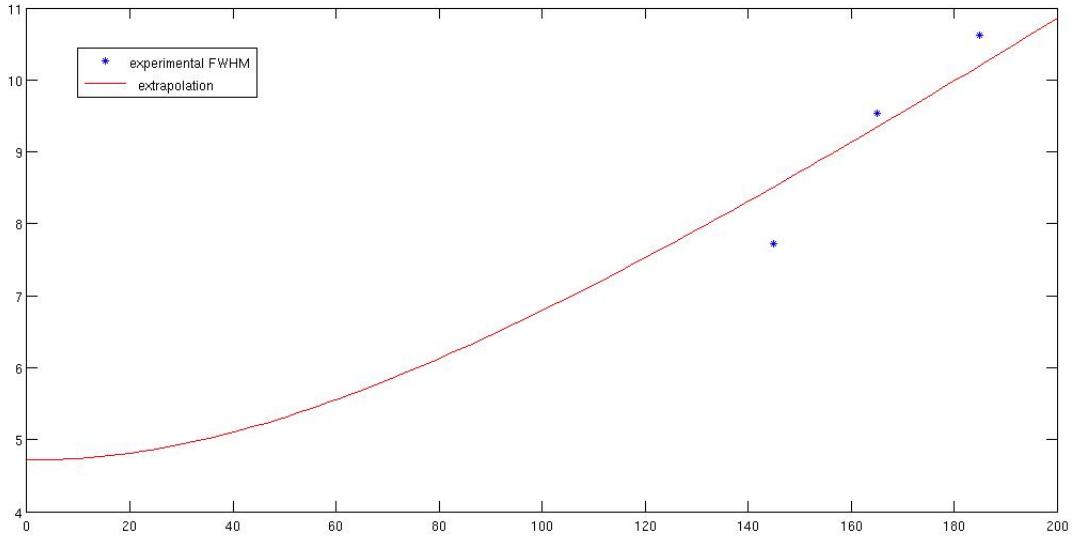


Figure 9.3: Extrapolation results from FWHM data using a nonlinear optimization function.

### 9.3 Cerebral dummy



Figure 9.4: The cerebral dummy used in experiments.

The second dummy used in experiments is a cerebral dummy like the one showed in Figure 9.4. This dummy is done with materials with the same attenuation coefficients of the human head. Instead of the brain is filled with water (the attenuation coefficients are similar) and tracer, while the four buffers on the inside are filled with a more concentrated tracer.

Again the raw data is a 128x128x120 matrix with 128 slices at different heights. In this case the data is very “poor”, because the matrix contains only integral numbers between 0 and 17, while the previous dummy data were up to 39. The low variance of the data is an index of a great attenuation and low concentration of the tracer<sup>2</sup>. So it is more difficult to reconstruct this data than the numerical phantom. As said before, we will only show some reconstructions and comment them.

---

<sup>2</sup>the tracer concentration and activity are the same used in real analysis on patients, so it's very low for security reasons.

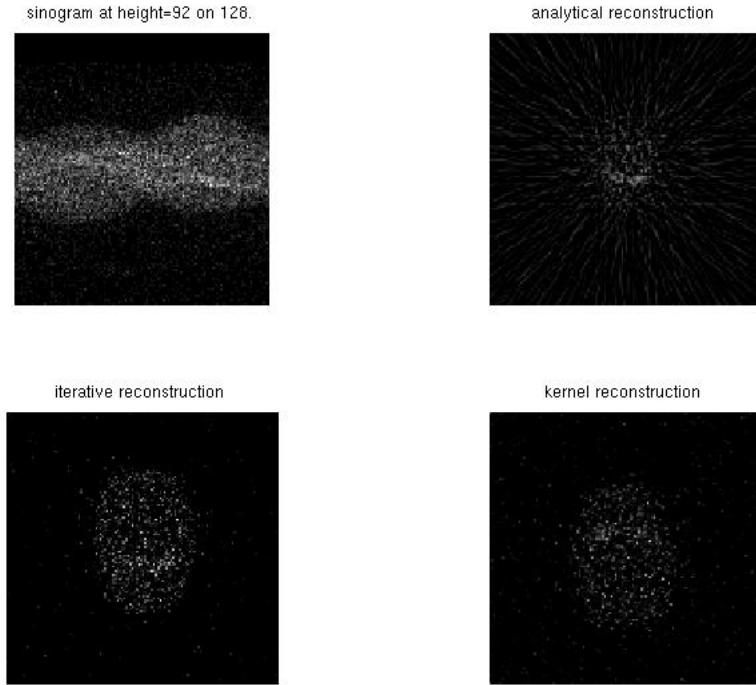


Figure 9.5: Different reconstructions of the sinogram data at slice number 92.

We have selected one slice and we have made different reconstructions, shown in Figure 9.5, and we have put all the data less than a chosen  $\delta$  to zero to eliminate some noise. The first reconstruction is made with FBP,  $\delta = 4$ . We have also added some zeros to the data for reaching the standards dimensions of the algorithm. As we see from the Figure, we can recognize the buffers, but the shape of the dummy is not clear. The second reconstruction method is an iterative MLEM after 100 iterations. We can identify both the buffers and the dummy shape in this image. The last method is a kernel interpolation using the Gaussian function with  $\varepsilon = 1.5$  and an MLEM at 100 iterations. In this case we can recognize clearly the shape of the dummy, but the buffers on the inside are disappeared.



# Conclusions and future work

Briefly, by points, we will summarize the work done and what remains to do.

## Work done

- Introduced CT and SPECT tomographies both from a medical and from a physical point of view (see Ch. 1).
- Studied the state-of-art algorithms for CT and SPECT (see Ch. 3-4).
- Found an error bound for the FBP formula (see Ch. 3.4).
- Introduced Kernel Methods with a relative error bound (see Ch. 5).
- Proposed an iterative method for SPECT/CT (See Ch. 6.4).
- Studied some methods for the regularization (see Ch. 7).
- Made some *In silico* and *in vitro* experiments (See Ch 8-9).
- Studied the connection between the resolution of the machine, of the sampling and experimental (see Ch 1.6, 3.2, 9.1).
- Proposed a method of extrapolation of the experimental resolution at a given distance from some data (see Ch 9.2).

## To do:

- Find an error bound for SPECT/CT algorithms
- Implement OSEM algorithm and test it with Kernel methods
- Study the setting of  $\lambda$  in iterative SPECT/CT
- Find an algorithm for contour detection in “poor-data” case

- Further *in vitro* experiments for the Kernel methods
- Extend the Kernel methods for 3D reconstructions

# Bibliography

- [1] Chang L.T., *A method for attenuation correction in radionuclide computed tomography*, IEEE Trans Nucl Sci Vol NS-25 No. 1, pp 638 - 643, Feb 1978
- [2] Cherry S.M., Sorenson J., Phelps M., *Physics in Nuclear Medicine*, Saunders, 2003
- [3] De Marchi S., Iske A., Sironi A., *Algebraic Medical Image Reconstruction from Scattered Radon Data by Positive Definite Kernels*, submitted, 2012
- [4] Fasshauer G.E., *Meshfree Approximation Methods With MATLAB<sup>®</sup>*, World Scientific, 2007
- [5] Feeman, *The mathematics of medical imaging: A beginners guide*, Springer, 2010
- [6] Freeman and Johnson's *Clinical Radionuclide Imaging*, Vol3: *Update*, Grune & Stratton, 1986
- [7] Gullberg, G.T., Yu-Lung Hsieh, Zeng G.L., *An iterative algorithm using a natural pixel representation of the attenuated radon transform*, Nuclear Science Symposium and Medical Imaging Conference, 1994
- [8] Hansen P.C., Hansen M., *AIR Tools - A MATLAB<sup>®</sup> Package of Algebraic Iterative Reconstruction Methods*, Journal of Computational and Applied Mathematics, 2011, doi:10.1016/j.cam.2011.09.039
- [9] Helgason S., *Radon Transform Second Edition*, Birkhauser Boston, 1999, [www-math.mit.edu/~helgason/](http://www-math.mit.edu/~helgason/)
- [10] Harbert J.C., Eckelman W.C., Neumann R.D., *Nuclear medicine; Radionuclide Imaging, Radioisotopes, therapeutic use*, Thieme Medical Publishers, 1996

- [11] Kak A.C., Slaney M., *Principles of Computerized Tomographic Imaging*, Society of Industrial and Applied Mathematics, 2001, [www.slaney.org/pct/pct-toc.html](http://www.slaney.org/pct/pct-toc.html)
- [12] Natterer F., *Inversion of the attenuated Radon transform*, Institut fur Numerische und instrumentelle Mathematik, 2000, [wwwmath1.uni-muenster.de/num/Preprints/2000/natterer/paper.pdf](http://wwwmath1.uni-muenster.de/num/Preprints/2000/natterer/paper.pdf)
- [13] Natterer F., *The mathematics of computerized tomography*, SIAM: Society for Industrial and Applied Mathematic, 2001
- [14] Novikov R.G., *An inversion formula for the attenuated X-ray transform*, preprint, Département de Mathématique, Nantes, 2000
- [15] Schaback R., Holger Wendland H., *Characterization and construction of radial basis functions*, in Multivariate Approximation and Applications, Cambridge University Press, 2001
- [16] Sironi A., *Medical Image Reconstruction Using Kernel Based Methods*, Master's Thesis, University of Padova, 2011, [arxiv.org/pdf/1111.5844v1.pdf](http://arxiv.org/pdf/1111.5844v1.pdf)
- [17] Soreson J.A., Phelps M.E., *Physics in nuclear medicine*, Second edition, Grune & Stratton, 1987
- [18] Toft P., *The Radon transform: theory and implementation*, Ph.D. thesis. Department of Mathematical Modelling, Technical University of Denmark, 1996, [petertoft.dk/PhD/](http://petertoft.dk/PhD/)